Security perspectives on social media exploitation

Nora Bieteniece

NATO Strategic Communications Centre of Excellence, Riga
LATVIA
nora.bieteniece@stratcomcoe.org

Valentina Dragos

Chemin de la Hunière, 91123 Cedex Palaiseau
FRANCE

valentina.dragos@onera.fr

Bruce Forrester

Defence Research and Development Canada, Quebec City
CANADA

bruce.forrester@drdc-rddc.gc.ca

Tomas Krilavičius

Vytauto Didziojo Universitetas Kaunas, Baltic Institute of Advanced Technology, Vilnius
LITHUANIA

tomas.krilavicius@bpti.eu

Albert Pritzgau

Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Wachtberg
GERMANY

albert.pritzkau@fkie.fraunhofer.de

ABSTRACT

Social media exploration for defence and security applications presents challenges that are beyond the ability of one domain or discipline to address. From online propaganda and disinformation to influence campaigns, social networks and platforms are often the vectors of intentionally distorted and biased narratives. Specific tools, like virtual bots, are used to manipulate public opinion and perceptions by amplifying or repressing political content, disinformation operations, hate and violent speech, and junk or fake news. Effective solutions for cyberspace exploration require synergies from organizational sociology, human computer interaction, communication, information science, and political science to interpret and analyze the evidence. The goal of this paper is to understand how social media streams and mechanisms can be better exploited from defence and security-driven perspectives. The paper discusses main challenges of social media exploration and relevant methods and techniques for use by military and civilian analysts in order to provide situational awareness and intelligence. Practical illustrations and use cases are also presented, explaining the possibilities and potential impact of the exploitation of social media channels for security purposes. The paper concludes by highlighting open questions and sketching future developments in the field.

1.0 INTRODUCTION

The Internet is currently a major source of information used by organizations and citizens to become informed on various subject and situations. The lack of social accountability in many digital platforms yields plenty of incentives for unprecedented forms of misuse. Disinformation, propaganda and fake news are just a few examples of ill-uses lurking in this largely accessible technology.
This paper aims to sample the state of the art in using social media analytics for security applications. First, we provide an overview of main challenges of social streams exploration and discuss methods developed to overcome those challenges. As the Internet is dynamic and heterogeneous, methods are based on hybrid Artificial Intelligence techniques. Secondly, we discuss several use cases of social media exploitation for defense and security applications. The main goal of the paper is to provide technical and practical insights in order to understand current capabilities for social streams exploration for security applications.

2.0 CHALLENGES OF SOCIAL MEDIA EXPLOITATION

Analysis of online streams comes with challenges stemming from both the specific nature of data created on digital platforms and the difficulties of social media exploration, as discussed hereafter.

Specificities of data: data collected from social media are vast, noisy, unstructured, inherently dynamic and heterogeneous in nature. Moreover, they convey reports on real-life facts and events augmented with personal points of view, such as evaluations, attitudes, and emotions. The data is, for the most part, inherently social in nature and reflects the inconsistencies and uniqueness found within humans. Our brains have evolved to be able to de-conflict and disambiguate the natural language that is found within social media posts. However, our ability to program algorithms to do the same is still evolving. Therefore, social data analysis is challenging for traditional data mining approaches that are often too slow and expensive, rely on sample sizes, and come with biases leading to errors.

Limitation of access to data and impact of secondary sources: An important volume of online data is released on DarkWeb [4], a problematic side of Web made of encrypted portions of the Internet that are not indexed by search engines and thus cannot be listed on results pages returned by search engines to user queries. Collection, processing and sharing of such content require specific procedures to be set up or the use of secondary sources. Nevertheless, secondary sources can introduce biases, or even truncate or alter the original set.

Influence of platforms and media-induced bias: Cyberspace is an artificial, manmade environment, with data and interactions framed in a particular manner. For this reason, social media platforms induce bias in how information is viewed by observers. Moreover, those observers can either be a part of the platform when undertaking their analysis or adopting a more direct approach to collect data via technical procedures. Those procedures are built on application programming interfaces (API), subroutines provided by platforms to access their collections of data, which are also stored in proprietary formats. This is a major drawback, as efforts are required to translate data into formats easier to process.

Security and ethical constraints, privacy protection: There are significant security related and ethical constraints to obtain first-hand information about sites, portals or content on social media platforms created by terrorists or extremist groups, from intelligences services, for example. Regarding privacy, data gathering and remote analysis for research purposes requires procedures and techniques that should be employed lawfully, as to make sure the overall process stay within the law.
For example, some countries within NATO do not allow their military intelligence analysis to collect and analyze data on their own citizens. This fact has allowed foreign agents to pose as citizens of these countries by using fake accounts. During the 2016 US presidential election, “Facebook believes 120 fake

Russian-backed pages created 80,000 posts that were received by 29 million Americans directly, but reached a much bigger audience by users sharing, liking and following the posts" [12]. This ploy is akin to an 'insider threat' and is very hard to detect.

In additional to those general challenges, social media exploration is also affected by technical bottlenecks, such as: multilingual issues, multimodality content, relevance and coherence of data sets, contextual information, aggregation and correlation of items, etc.. Next section discusses methods and techniques developed to address some of those challenges.

3.0 METHODS AND TOOLS FOR SOCIAL MEDIA EXPLORATION

Effective and efficient intelligence is usually accomplished in a combined human and computer effort. Indeed, the intelligence process heavily depends on combining the human's flexibility, creativity, and cognitive abilities with the capacity and processing power of today's computers.

Exploitation of social media generally breaks down into three fundamental objectives: information discovery, situational awareness and predictive analysis. Capabilities in addressing these objectives provide essential estimates of the potential risks faced by communities, economies and the environment. When exploiting social media sources, analysis is commonly limited to two aspects: users as basic units of the network and content as basic elements of communication [19]. These two aspects themselves are already invaluable sources of information. However, social networks additionally offer the context of communication and interaction represented by the network itself, namely, the network topology in the form of entities and relations. In addition to content, a given network structure promotes the derivation of activity and process patterns which can significantly improve situational awareness [20]. As a result, different data representations rendering distinct aspects of content and interactions serve as a means to adapt the focus of the intelligence analysis to specific information requests.

Methods and techniques for social media exploration can be roughly divided into lexicon-based [1] and machine-learning methods [3], and analyse the implicit, explicit and discursive patterns within data sets [2]. More specifically, content analysis and natural language processing (NLP) are two techniques of particular interest for social media exploration.

Content analysis is the quantitative analysis of properties inherent to different forms of communication. It is a scientific method used to systematically and reliably quantify the symbols used in communication. Content analysis has come to be seen as a particularly useful method for analyzing communication techniques associated with computerization and digitization. Indeed, user-generated contents have several interesting properties such as diversity, coverage and popularity that can be used as wisdom of crowds in search processes. Methodical approaches to content analysis, text analysis and text mining examine artefacts of social communication, typically written documents or transcripts, to develop objective inferences about a subject or topic of interest [21], the framing of conversations and many other characteristics of messages. Considering the huge volumes of social media data being created, it becomes clear that automated methods of text analysis make an invaluable contribution.

Promising ways to efficiently extracting intelligence information are NLP methods. These methods usually try to reframe complex language understanding tasks as simpler classification problems. Today's models produce impressive results on some of the challenging tasks including Question Answering, Sentiment Analysis, and Text Entailment. Since information campaigns and propaganda also use psychological and rhetorical techniques to reach purposes, tools are developed to detect automatically these techniques by reframing them as classification tasks. Of particular interest to intelligence is to understand the collective behavior of certain groups of people, which is often supported by narratives. By automatically categorizing messages into different narratives about events and topics, we aim at raising the analyst's awareness into the emergent views. The resulting analytical models can be used to reduce huge amounts of data to relevant data streams. The resulting subsets can then be subjected to further investigations such as the detection of propagandistic elements.

From practical perspectives, applications of social media analysis aim at: online hate [5] and violence detection [6], extremist content analysis [7], identification of online communities [10, 13], opinion leaders

[7, 13] and contagion mechanisms [8], analysis of subjective content [11, 13] and investigation of online behaviour [9].

An approach called authentic chatter [13] was developed to help overcome the problems of large data sets and the need to have a good understanding of both influencers and narratives within topic areas. This Twitter specific method exploits social network analysis research and qualitative analysis. Relevant topic areas are identified by requirements. The prominent influencers (authors) within these areas are determined by using indegree and retweet metrics. This prominent group is then qualitatively analysed and assigned a type (official, news, SPAM, BOT, or Authentic). In this way the posts of between 50 – 100 authors can be analysed in detail in order to provide an effective proxy for the entire topic area. This method significantly reduces the time required for topic awareness while an in-depth understanding of the influencers and narratives is gained by the analyst. As such, appropriate responses and actions can be recommended.

The role of exploratory search for sensemaking in the intelligence process has already been described by Gersh et al. [14]. Exploratory search can be characterized by successive queries interspersed with stages of sensemaking [15]. According to Marchionini, general search activities can be grouped in "lookup", "learn", and "investigation" [16]. In particular, Wilson et al. [17] characterized investigation and learning as the two most important goals in exploratory information seeking, both of which require the involvement of the analyst in a range of activities such as comparison, aggregation, and evaluation. To this end, advanced information technologies and tools are developed to improve the efficiency and accuracy of intelligence analyst's work. Usually, a query is the bridge between the analyst's conception of the prevailing information need and the information access system [18]. To close simple information gaps, a query specification accomplished using keywords, key phrases, Boolean operators combined with command-based syntax may be sufficient. Complex processes of analytical reasoning, however, call for more elaborated mechanisms to specify queries tuned to a specific research question.

As highlighted by the selection of research efforts above, social media exploration is a challenging task requiring the development and training of domain specific analytical models, the incorporation of diverse knowledge resources- the analyst's expertise, ontologies, knowledge graphs and more- and the implementation of hybrid methods relying on both data and knowledge driven techniques.

4.0 DEFENSE AND SECURITY USE CASES

This section illustrates several use cases conducted to explore social media for defence and security applications.

4.1 Propaganda detection

Case study: This case study aimed at detecting Russian influence. Two simple filters were developed in order to detect suspected Russian-based tweets. One filter looks at content and the other examines users. The first filter consists of a list of 200 websites taken from www.propornot.com. This site has, and continues to, identify sites that produce or propagate Russian propaganda. The second filter identifies authors who are associated with the Internet Research Agency (IRA) and who have been labelled as Russian Trolls by Twitter. This data set was released by Twitter in 2018.

Application context: These filters were recently applied to a Twitter data collection that had as an aim to detect foreign influence campaigns during the lead up to the 43rd Canadian election in October 2019. Past influence efforts, the 2016 US presidential election, usually involve trying to divide the opinions of the population. The campaign would focus on wedge issues, where there were opposing views. Efforts concentrated on moving one or both sides to a more radical view thus creating a seemingly irreconcilable divide between people.

To do this well-designed BOTs are employed. The BOTs are assigned a certain view – usually one that is in the interest of the foreign influencer. However, sometimes the aim is just to divide the people such that a minority (weaker) government is elected. Once the BOT focus is decided, one technique used is that the

BOTs are employed to support the chosen side using relevant hashtags and linking to local websites that push extreme views for a given topic. This technique makes it seem that there is a lot of support for the views within the extreme website at a local level thus encouraging others to adopt these views.

Methodology: For this case, we first identified potential contentious issues that were likely to be discussed during the election. A data collection, using appropriate hashtags and keywords, was started for each issue several months preceding the election. Care was taken to isolate the issues and ensure that mainly Canadian users were captured. Note that we were not looking at individual Canadians; the focus was on finding foreign BOTs within topic areas. For each issue we applied the two filters to determine if there was links with Russian propaganda or trolls. The first issue we examined was a smaller data set and neither filter produced results. However we were able to identify several BOTs within this topic. Next we examined a much larger topic and the Russian propaganda filter produced results. We examined the users identified by the filter and found a large proportion of BOTs. At this point we were able to use the BOT handles from these two topics as seeds to quickly examine the other wedge issues.

Analysis of results: We uncovered a large BOTNET that had many BOTs involved in several issues and at least two BOTs that posted in all of the wedge issues. Applying the IRA filter did not produce any results. We believe that, since the Russian Troll handles were compromised by their release by Twitter, these handles are no longer in use. A possible replacement for the user-based filter would be to use the BOT handles. Overall, this method proved very effective by producing positive results in less than 3 hours of effort.

4.2 Opinion detection in social streams

Case study: This study aimed at detecting opinions in social media and investigated the use of appraisal categories to explore social data [22]. The approach was grounded on cognitive foundations of the appraisal theory, developed by White and Martin [23]. The appraisal theory introduces three cognitive systems making the distinction between affect, appreciation or judgement, and highlights the mechanisms by which text convey positive and negative attitudes. The approach goes beyond the generally accepted definitions of sentiment and opinions, and focuses on appraisal expressions to describe the way humans express their attitudes, appreciations, and engagement.

Application context: The application context of this work is a cyberspace exploration task designed to support defence and homeland security intelligence practitioners in their efforts to gather valuable data allowing them to understand emergent phenomena, such as online hate proliferation or online propaganda. Linguistic clues of appraisal categories are used as indicators of subjective content to be collected and further analysed in order to understand the way online users express their extreme attitudes, embrace or support extreme ideologies and ideas.

Methodology: Starting with the three systems introduced by the appraisal theory, a semantic resource was build modelling finer categories under each system, and highlighting terms and expressions specific to each category. The methodology relies on a semantic annotation approach augmented with processing methods capable of performing a quantitative analysis of data gleaned on social media. The approach consists of several phases implemented to gather, process and analyse social data.

- Data acquisition: was done by crawling several sets of tweets with specific keywords and additional constraints to select only posts written in English, see tab. 1. This phase was carried out as a straightforward step and relies only on data content of both keywords and posts and the ability of API used to mine the social network.
- Data processing: performs first a cleaning step, by removing URLs, hashtags and any information considered as irrelevant. Processing at paragraph level includes sentences identification according to punctuation marks, tokenizing, part-of-speech (POS) tagging along with identification of words stems.
- Semantic annotation: attaches additional information to various text paragraphs based on their content analysis.
- Data analysis: was carried out based on semantic annotations and estimates for each set of data; the percentage of subjective versus objective tweets; the percentage of tweets having positive

versus negative orientations; the distribution of tweets according to their low, medium or high strength and also their distributions with respect to *Attitude* and *Engagement* systems, and also with a finer distribution according to their specific concepts.

Table 1: Data collections for opinion analysis

Key word	Remarks	Number of tweets
Western values	Often associated with online hate	7956
White supremacy	Extremist content	7542
Iraq war	Old conflict	7001
Western coalition	Often associated with online hate	6971
Security	Generic term	7509

Results: The analysis of data sets was carried out in order to calculate numerical distribution of subjective and objective tweets, positive and negative orientation, high and low force along with a finer analysis in the light of appraisal categories. Numerical values show a high percentage of subjective tweets for all datasets analysed, regardless of the keyword. Tweets conveying *Attitude* concepts are generally less than 50%. There is a good representation of positive-oriented tweets that account for more than 50% of the collection, while tweets having low or high impact are underrepresented for all data sets. The method also highlights associations of key words and concepts: for example *White supremacy* has strong correlations with *valuation* and *affirm*, and low connections with *impact* and *concede*. At collection level, the analysis shows a rather similar distribution of data sets into categories of appraisal systems, with around 10 classes for *Attitude* and 7 or 8 classes for *Engagement*.

4.3 Dynamics of Lithuanian radical groups in Facebook

Data collection: Data was collected from Facebook using FG graph API [1] (currently not available) on February 12, 2018. Posts from period March 4, 2010 – January 1, 2018 were downloaded. Posts, post creation data and posts ids were collected and stored. The groups were identified using Facebook search engine, mostly based on posts proclaiming nationalism, strong nation, xenophobic ideas as well as hostility to ethnical and sexual minorities, i.e. the following keywords were used *Lithuania, Lithuanians, land sale, European Union, NATO, refugees, refugee crisis, Muslims, Jewish restitution, Jew, Russian, Roma tabor, gay pride, gay mountaineering.*

After preliminary analysis, 10 groups were chosen for the in-depth analysis, based on the following criteria: (1) presence of the radical ideology, (2) more than 100 members and (3) the last post was published at least 2 days prior to the analysis (i.e., group is still active).

Two datasets were collected, see tab.2:

- Pro-Russian FB groups: 70 150 posts, 13940 members.
- Other radical right groups: 9578 posts, 6126 members.
- Total: 79728 posts.

Analysis methods: Most frequent words were used as features, and posts timestamps were used to identify change of activity during the all period. Texts were not lemmatized.

Table 2: Data collections for dynamics analysis

	Year							
	2010	2011	2012	2013	2014	2015	2016	2017
Pro-Russian	0	0	0	0	10447	17033	19257	23413
Radical right	58	311	604	1164	2997	791	1914	1739
Total	58	311	604	1164	13444	17824	21171	25152

Analysis of results: Activity (see fig. 1) of two types of groups was quite different:

- Pro-Russian groups were active in FB from 2014, and the peak was 2017.
- Radical right groups were active from 2010, activity growth was very fast, and reached peak in 2014.

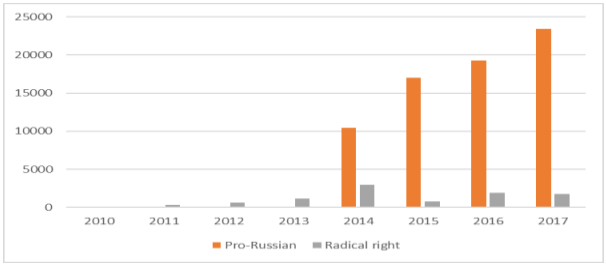

Figure 1: Activity of FB groups

Moreover, appearance of different keywords in posts was analysed, e.g. Lithuania was mostly mentioned in 2014 by pro-Russian groups, and 2017-2018 by other groups, but it was quite important all the time, while Russia is mostly discussed by pro-Russian groups (with peak in 2015) and only scarcely by other groups, mostly in relation to sanctions and Ukrainian crisis.
Basically, the results show that FB groups' activity is strongly influenced by different events, involving Lithuania and Russia, and most of radical groups are moving to Facebook from forums and other media channels. Monitoring of such groups could be a good indicator for different groups' activities, dynamics of such groups, and related risks.

4.3 Target identification

Target identification, i.e., identification of domains or information actors in the online environment, is an important step in social media exploration for defense and security. It narrows in on outlets or online actors that exhibit behaviour of interest or belong to a real-life group of interest. It also reduces the amount of information to be processed when exploring social media.
Open source intelligence (OSINT)[1] methods and tools are often used for target identification. This includes manual monitoring and processing of social media streams, user data and metadata or the use of automated OSINT tools[2] for monitoring and processing such information. Both approaches can be generally divided into text-based methods/tools, geographic information and remote-sensing

[1]The process of visual and textual monitoring of and deriving information from publicly available data in the intelligence context is a common.
[2] Maltego, TinEye, IntelTechniques.

methods/tools, network science methods/tools and visual forensics[3] none of which are discussed in this paragraph.

Application context and methods: Method frequently used in conjunction with OSINT tools and techniques is crowdsourcing, the practice of obtaining information or input by enlisting the services of a large number of people via the Internet. Much of Bellingcats'[4] work is an example of crowdsourced target identification. In 2014, Belligcats' crowdsourced analysts used open-source tools to discover which Russian unit shot down the MH17 flight in Ukraine[5]. For this analysis Bellingcat volunteers used images and social media posts taken and posted by the local public prior and/or after the accident. They also used satellite imagery to verify the locations identified from social media posts and images, and searched and found their suspect actors online. Their analysis was sufficient to be used as evidence by Dutch Safety Board[6].

A less commonly known target audience identification method in conventional defense and security is honey-potting. In cybersecurity honeypot is a closely monitored machine serving several purposes: it can distract adversaries from more valuable machines on a network, provide early warning about new attack and exploitation trends, or allow in-depth examination of adversaries during and after exploitation of a honeypot. In the application of social media exploration for defense and security, honey-potting is a space designed to "lure" possible suspects or targets. In NATO StratCom COE recent study[7], Facebook page run by the read-team was launched and advertised to possible military exercise participants via Facebook Ads. After gathering a significant following, the red-team created a closed group for exercise participants and advertised it via the said page. This group served as a "honey-pot" for identifying exercise participants; moreover, it provided a space to interact with the participants without being challenged by the general public. This method of target identification is more "invasive" in comparison with OSINT techniques that merely gather, correlate, verify and create new evidence from existing data points such as social media posts, images etc. However, it serves potentially as a source for new information that suspects or targets otherwise would not have disclosed on their social media feed.

The variety of use cases discussed above clearly illustrates the potential of social media to facilitate the creating and sharing of information useful for defence and security applications. However, those use cases also emphasize the crucial importance of social, linguistic and cultural indicators, the difficulty of developing efficient approaches for social media exploitation and the exploratory nature of the task.

5.0 CONCLUSION AND PERSPECTIVES

This paper provides an overview of challenges and methods for social stream analysis. This an important and timely endeavour because social media does not only play a significant role in shaping people's attitudes and beliefs, but it also has far-reaching consequences for societies in general, such as increasing tendency to violating social norms and disrupting democratic processes. The paper analyses several approaches used for social media analysis. The first was based on filtering for content or users that are known to be used by Russia. The second and third case studies use linguistic cues that are based on sociology principles and aim at understanding human behaviour online and the propagation of stories of interest to various target audiences. The last use case discusses the identification of targets on social networks and the analysis of interactions between agents in the real and virtual environments for intelligence tasks.

As a dynamic and heterogeneous environment, social media exploration deserves further study and leads to interesting new perspectives when combined with relevant research in related areas such as information

[3] https://edam.org.tr/wp-content/uploads/2018/07/Digital-Open-Source-Intelligence-Bosch-2nd-Report.pdf
[4] https://www.bellingcat.com/
[5] Ibid.
[6] https://www.bellingcat.com/news/uk-and-europe/2015/10/15/how-the-dutch-safety-board-proved-russia-faked-mh17-evidence/
[7] https://www.stratcomcoe.org/current-digital-arena-and-its-risks-serving-military-personnel

retrieval, event detection or trust analysis. Certainly the major improvements in hardware power and categorization algorithms have led to significant advancements in the research. However, as researchers and analysts become better at detection of influence and deception, those wishing to influence and deceive are continuously creating new techniques making it a 'cat and mouse' game. Challenges still exist. Natural language processing, sentiment analysis and content analysis continue to be problematic when applied to the messy nature of social media data. Further, the always increasing and never ending stream of data taxes our analytical resources. We must continue our efforts to increase the signal to noise ratio allowing analysts to concentrate on pertinent signals. Finally, our analytical tools need to be precise thus permitting for reliable, valid, and accurate assessments that can translate into actionable intelligence for decision makers.

6.0 REFERENCES

[1] M. Taboada, J., Brooke, M., Tofiloski, K., Voll, and M., Stede. Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), 267-307., 2011.

[2] F., Villarroel Ordenes, S., Ludwig, K, De Ruyter, D., Grewal, & M., Wetzels, (2017). Unveiling what is written in the stars: Analyzing explicit, implicit, and discourse patterns of sentiment in social media. *Journal of Consumer Research*, *43*(6), 875-894.

[3] E., Cambria, C., Havasi, and A. Hussain, SenticNet 2: A Semantic and Affective Resource for Opinion Mining and Sentiment Analysis. In *FLAIRS, 2012.*

[4] Weimann, G. (2016). Going dark: Terrorism on the dark web. *Studies in Conflict & Terrorism*, *39*(3), 195-206.

[5] Isbister, T., Sahlgren, M., Kaati, L., Obaidi, M., & Akrami, N. (2018). Monitoring Targeted Hate in Online Environments. *arXiv preprint arXiv:1803.04757*.

[6] Alvari, H., Sarkar, S., & Shakarian, P. (2019). Detection of Violent Extremists in Social Media. *arXiv preprint arXiv:1902.01577*.

[7] Wei, Y., & Singh, L. (2017, May). Using network flows to identify users sharing extremist content on social media. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 330-342). Springer, Cham

[8] Ferrara, E. (2017). Contagion dynamics of extremist propaganda in social networks. *Information Sciences*, *418*, 1-12

[9] Lara-Cabrera, R., Gonzalez-Pardo, A., Barhamgi, M., & Camacho, D. (2017, August). Extracting radicalisation behavioural patterns from social network data. In *Database and Expert Systems Applications (DEXA), 2017 28th International Workshop on* (pp. 6-10). IEEE.

[10] Rossetti, G., Pappalardo, L., Pedreschi, D., & Giannotti, F. (2017). Tiles: an online algorithm for community discovery in dynamic social networks. *Machine Learning*, *106*(8), 1213-1241.

[11] Ruhrberg, S .D.,Kirstein, G., Habermann, T., Nikolic, J. and Stock, W.G.(2018) #ISIS—A Comparative Analysis of Country-Specific Sentiment on Twitter. Open Journal of Social Sciences, 6, 142-158.

[12] Solon, O., & Siddiqui, S. (2017). Russia-backed Facebook posts 'reached 126m Americans' during US election. The Guardian. 31 October, 2017. Accessed 9 September 2019.

[13] Forrester, B., (In Press) Authentic Chatter, Computational and Mathematical Organization Theory Journal.

[14] Gersh, J., Lewis, B., Montemayor, J., Piatko, C., & Turner, R. (2006). Supporting insight-based information exploration in intelligence analysis. *Communications of the ACM*, *49*(4), 63-68.

[15] P. Pirolli and S. Card, "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis," *Proc. Int. Conf. Intell. Anal.*, vol. 2005, pp. 2–4, 2005.

[16] G. Marchionini, "Exploratory search: from finding to understanding," *Commun. ACM*, 2006.

[17] L. Wilson, "From Keyword Search to Exploration: Designing Future Search Interfaces for the Web," *Found. Trends® Web Sci.*, vol. 2, no. 1, pp. 1–97, Jan. 2010.

[18] M. Hearst, "Search user interfaces," 2009.

[19] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?," in *Proceedings of the 19th international conference on World wide web - WWW '10*, 2010, p. 591.
[20] D. Helbing *et al.*, "How to Save Human Lives with Complexity Science," p. 67, Feb. 2014.
[21] K. H. Krippendorff, Content Analysis: An Introduction to Its Methodology, 3rd ed. 2012.
[22] Rasa Kasperienė, Tomas Krilavičius. Application of Social Network & Content Analysis Methods for Assessing the Dynamics of Facebook Groups. The 24th International Conference on Information Technology (IVUS 2019). Kaunas, Lithuania, Apr 25, 2019.
[23] Graph API in Facebook for Developers. https://developers.facebook.com/docs/graph-api, Accessed 19 Apr 2018