SafeGuard Cyber

CONTACTLESS ACTIONS AGAINST THE ENEMY

How Russia Is Deploying Misinformation on Social Media to Influence European Parliamentary Elections

TABLE OF CONTENTS

Executive Summary

The purpose of this report is to evaluate the behavior and activity of Russian misinformation campaigns conducted on social media in the run-up to European Union (EU) Parliament elections, from 23-26 May 2019. We focused on social media because it is ground zero for penetrating the consciousness and news cycle of any given EU member state. To that end, this report is a summary of SafeGuard Cyber's investigation into malicious information activity on social media, dating from November 2018 to March 2019.

The Russian misinformation campaign methodology is not hidden or new, but part of the publicly available so-called Gerasimov Doctrine. In an outline of a new approach to warfare, Russian Chief of the General Staff Valery Gerasimov proposes a vision of future military conflict in which "the information space opens wide asymmetrical possibilities for reducing the fighting potential of the enemy."[1]

Our goal was to understand if the tactics we observed in last year's Twitter bot report were still in place, how messages are disseminated, and to what extent EU government institutions or personnel are vulnerable.

For this undertaking we built new technologies and invested effort to improve existing detection capabilities to more efficiently study the massive amount of data involved. We segmented messages shared or promoted by bots, trolls, and other bad actors. From there, we classified content according to subject and scrutinized communication patterns for any common traits.

Our study found:

- Bad actors encompass three categories: bots, trolls, and hybrid bad actors
- Misinformation volume is directed at EU member states to exploit and exacerbate developing social fissures and contentious issues in near real-time
- Bad actors amplify content to either shape public perception of events, or project certain narratives into the zeitgeist at greater speed to bypass traditional media
- EU personnel across administrative bodies and rank are currently vulnerable to bad actor operations, putting EU digital infrastructure at risk

[1] https://usacac.army.mil/CAC2/MilitaryReview/Archives/English/MilitaryReview_20160228_art008.pdf

Methodology & Framework

For this report, our investigators used the SafeGuard Cyber platform to comb through almost 3.5 million posts on Twitter, Facebook, Instagram, and YouTube from November 2018 – March 2019.

We analyzed and evaluated tactics against the military strategy articulated in the Gerasimov Doctrine, principally the Russian perspective that "the role of nonmilitary means of achieving political and strategic goals has grown, and, in many cases, they have exceeded the power of force of weapons in their effectiveness."[2] For the Russian state, the information space is another arena of warfare, and social media represents the highest value information space, given its proximity to the local citizenry.

In this report, we will present a snapshot of 1-10 March 2019. In this particular time period, a number of real-world events converged in EU countries that put to light the clear orchestration of Russian misinformation tactics.

Bad Actor Nomenclature

In the technology we use to proactively protect our clients, we refer to malicious users broadly as "bad actors." However, there are some further distinctions to be made:

- Bots: Fully automated entities, usually run with scripts that allow them to post at volumes that would be impossible for a human. They may also be programmed to pick up specific hashtags or other text-based cues.
- Trolls: Human actors tasked with responding in a certain way or amplifying certain content. They may also disseminate content intended to be picked up by parallel bot battalions.
- Hybrids: Human actors using software to communicate through multiple accounts at the same time. This tactic may be used to avoid bot detection algorithms.

Our platform monitors and classifies all three types of bad actors. Where possible, we use the precise terms for each entity.

[2] https://usacac.army.mil/CAC2/MilitaryReview/Archives/English/MilitaryReview_20160228_art008.pdf

Technology & Sources

Our Threat Chain analytics engine uses machine-learning algorithms to assess social media accounts against 52 risk signatures which help us determine the probability with which a given account fits into four critical modes of behavior. We cross-reference all four categories with probabilistic analyses to arrive at an overall score indicating the likelihood an account is a bad actor:

1. **Malicious:** posts include malicious content or attempts to lure other users
2. **Suspicious:** posts exhibit characteristics that warrant further analysis
3. **Disinformation:** posts content known to be shared for misinformation purposes
4. **Bot:** posts at such a frequency and volume as to be artificial

Bad Bots vs. Benign Bots

Bots in and of themselves are not necessarily “bad.” A number of organizations and tools use bots to automate social posting for ease or to reduce human resource requirements. Many major news publications’ social accounts qualify as “bots” simply for the sheer volume of content being posted.

Example: an account scores high for bot activity and no other category, therefore it would not register as a bad actor. However, should it also register for malicious content or disinformation, then it would score as a bad actor.

To determine misinformation content, we built a proprietary tool that aggregates the data from 155 fact-checking sites (e.g., Politifact, EU vs Disinfo, Snopes. etc.) in 53 different languages, including English, French, Spanish, German, Russian, Georgian, Kazakh, Korean, and more. This fact-verification platform scans across social media, news, and forums, and cross-references misinformation content against our proprietary Bot & Trolls Database. We also use natural language processing and semantic analysis techniques to classify content. In this way, we can verify content and understand which bots and trolls are amplifying these stories. Our database contains over 500,000 known troll and bot accounts, and continues to grow. The database also allows us to isolate these bad actors and study tactics in closer detail.

Our Findings

Establishing Narratives

Misinformation agents work within clear narrative categories. Influence operations can appear difficult to discern because the content moving through conversations takes many forms and appears scattershot as any topic on social media. However, chasing any and every topic would actually dilute misinformation efforts, because campaign managers are aiming to achieve a certain "critical mass" of messaging in order to exert any influence on the average citizen. In this way, narrative constraints help maintain operational and messaging focus. Looking at content for this period, we saw similar categories as we have seen before. However, messaging was very much honed for a European audience.

Below is a table showing how these categories are defined in relation to their goals, including examples of events and conversations where bots or trolls were deployed.

Narrative Category*	Goal	Example
Sow Division	Destabilize, foment division	Brexit, "yellow vest" protests, euroskepticism
Revisionist History	Re-cast history in friendly light	Holocaust denial, distortion around founding of EU
Russian World	Tout Russian culture, government, etc.	"invincible" missiles, economic achievements, Eurovision
Russian Relations	Push friendly terms for Russia	NordStream II pipeline, anti-sanctions
Occupied Voices	Pro-Russian propaganda	Crimea, Georgia
Discredit	Undermine verified facts	Ukraine as "fascist" state
Conspiracy	Surface and legitimize fringe ideas	Soros influence, anti-semitic tropes

**This table represents a sample of the most frequently used categories of misinformation campaigns*

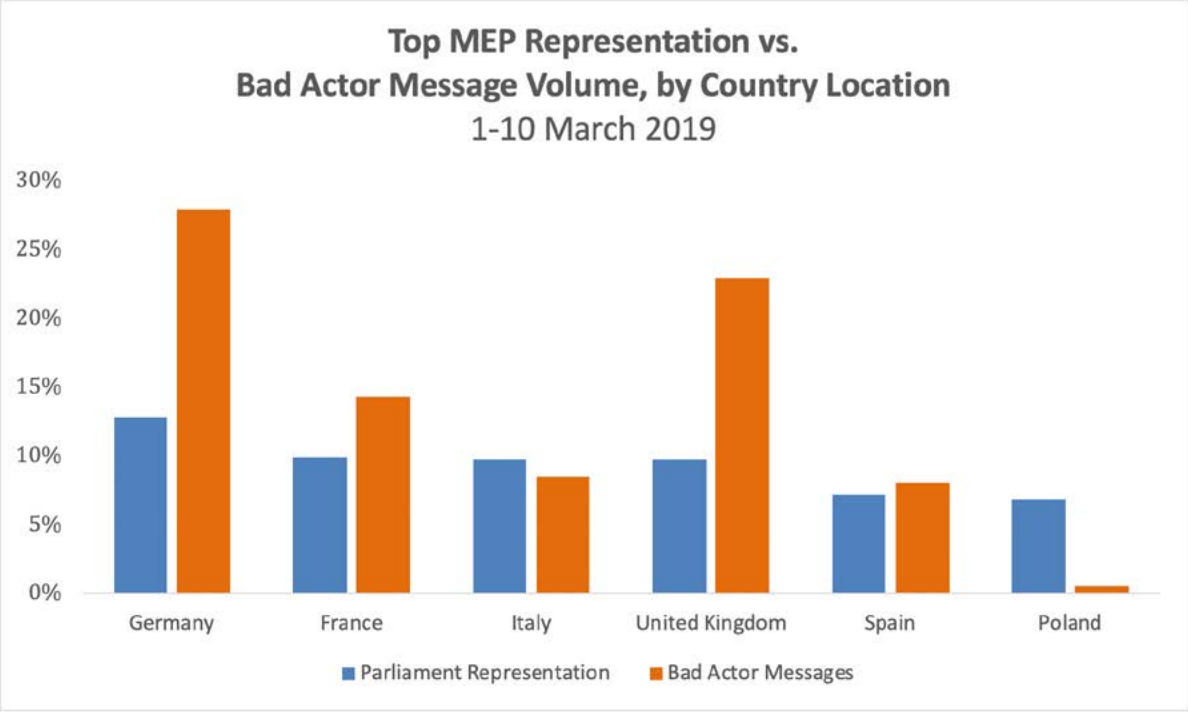
Since 2016, Russian operators have modified their tactics. In lieu of creating their own content from scratch, they now appear to spend more effort amplifying existing content that is being created organically. One might argue this development is a natural evolution. By opening divisions in society and empowering the extremes of society, Russia can simply rely upon those extremes to create the messaging. It's up to the bots to amplify extremist messages (right and left) to continue exacerbating divisions in society. However, it is also more tactically efficient to amplify existing content, and provides some cover from platforms' efforts to crackdown on fake accounts using AI and more sophisticated algorithms.

A Story for Every EU Country

There is no “all purpose” content being promoted to influence EU elections. Once again, operators are using bots, trolls, and hybrids to increase coverage and therefore give greater credence to existing content. To that end, there is a narrative attack crafted for every EU member state. Here we will highlight findings related to operations against the largest states, by parliamentary representation, and some outliers that bear closer examination.

Operations Against Larger EU Member States

At first glance, the level of influence operations against the larger members states appears to be commensurate with a target state’s level of diplomatic conflict with Moscow. Open critics appear to bear the brunt of misinformation campaign volume. The graph below appears to support this, with Germany, France, and the UK (Brexit notwithstanding) targeted with a disproportionate amount of bad actor-based messaging, relative to their MEP representation, which is a useful proxy for level of influence in overall EU affairs. This is a snapshot from 1-10 March, but is in line with what we saw for the entire time period of investigation.



Germany would be easy to explain with the aforementioned relationship. However, we believe that relying on an understanding of government-level relationships to explain misinformation volume is too simplistic an interpretation. Most importantly, this relationship ignores the goals outlined in the Gerasimov Doctrine of using “internal opposition to create a permanently operating front through the entire territory of the enemy state” (emphasis added).[3]

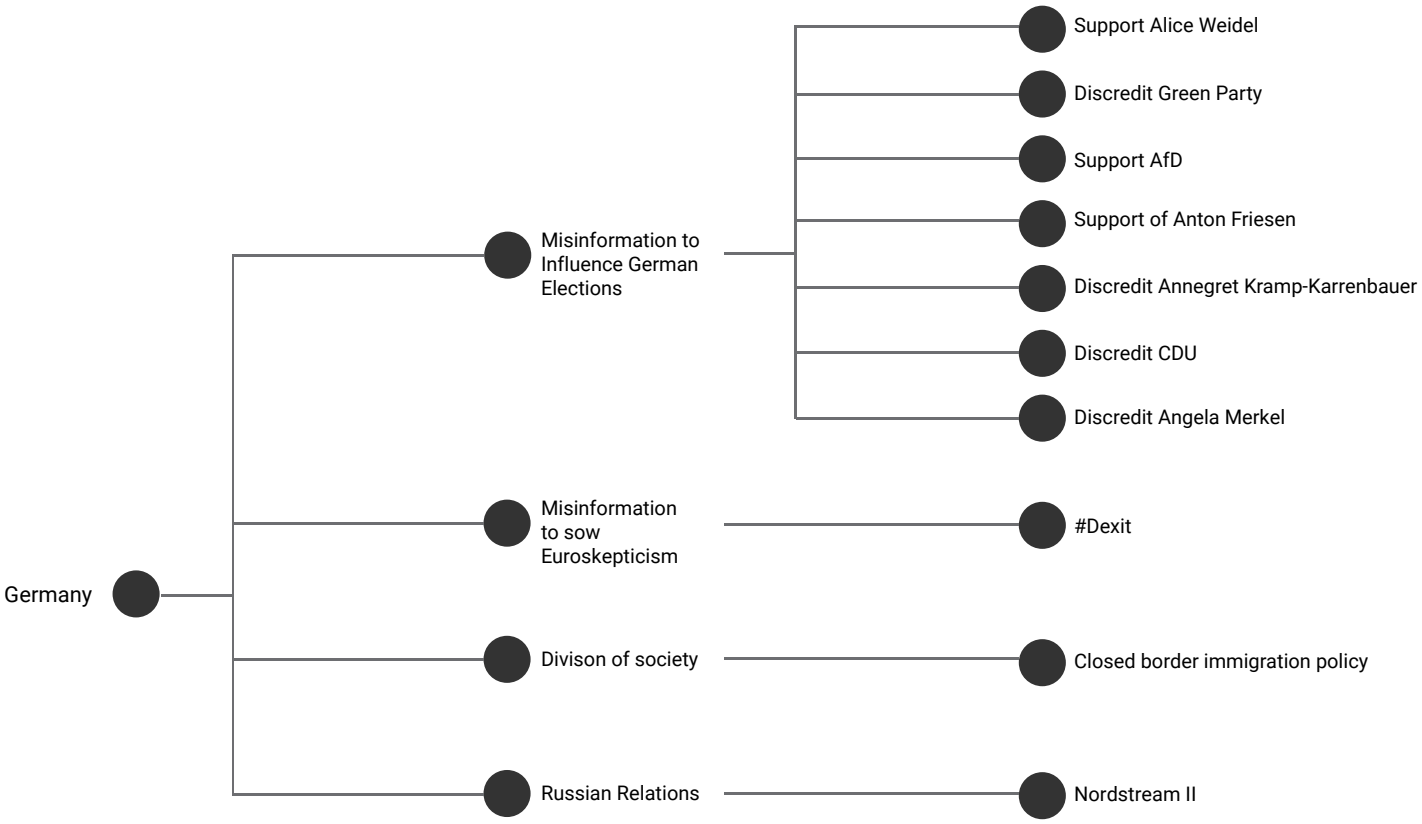
We propose that the effort leveled at a target state has more to do with the socio-political fault lines within said state that can be exploited. Where there is more high-stakes internal division, bad actors

[3] https://usacac.army.mil/CAC2/MilitaryReview/Archives/English/MilitaryReview_20160228_art008.pdf

will have an easier time of it. Misinformation operations are tied to larger Russian strategic goals, but these campaigns rely on a certain spirit of opportunism. Indeed, when we scrutinized the narrative structure of misinformation further we saw how existing conditions made these states (Germany, France, the UK) more susceptible targets.

Examining Country-Narratives More Closely

Applying the same natural language processing and semantic analysis techniques to target countries, we saw narrative categories break down into sub-narratives targeting clear friction points. In the case of Germany, this was content amplified to cause disruption in German elections and sow Euroskepticism. The former split into several themes, including supporting or discrediting specific high-profile politicians.



Below, one can see examples of posts amplified by bots and trolls against these themes.

Again, it must be reiterated that this content was not created by Russian misinformation operators. These posts are examples of the existing fissures in German society that are being inflated and given greater exposure than they might otherwise. As an example, an account that shared Udo Hemmelgarn's post appears to be reposting similar content at a rate of 2.3 posts per second for hours at a time.

In the post on the right, by a hybrid bad actor, the hashtags are disparate and wide-ranging. In a post ostensibly trying to blame Islam for a particular incident, the user is also trying to associate the content with

- #CDU (Christian Democratic Union, and party of Angela Merkel)
- #AKK (Annegret Kramp-Karrenbauer, Merkel's heir apparent)
- #Macron (French President Emmanuel Macron)

Moreover, this same series of hashtags is frequently appended to the account's posts. This tactic presents a two-fold danger:

1. The tags themselves are picked up by bots that automatically share or comment on anything with those tags, creating amplification nodes throughout social media.

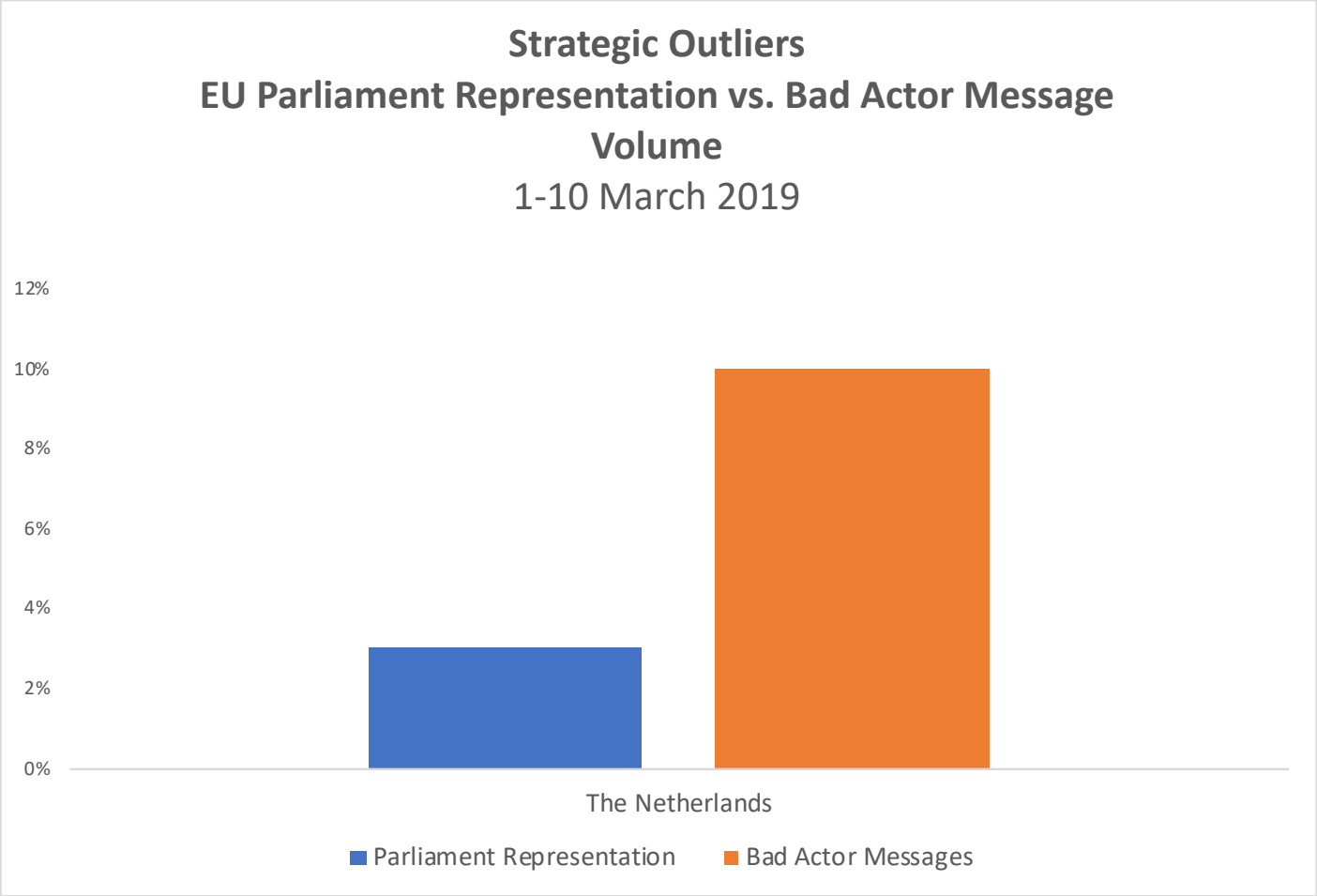
2. It is entirely possible and plausible that this sort of behavior will be picked up by real human users in target countries. Seeing what appears to be real posts that use hashtags this way might inspire or teach others that this is a "legitimate" behavior. Thus, real users will unwittingly be recruited into the amplification systems crafted by misinformation operators.

In summary, the states facing the highest volume of misinformation messaging are prime targets because of existing socio-political tensions. In Germany, this includes immigration policy in the wake of the Syrian refugee crisis and the rise of the AfD party; in France the "yellow vest" protests, and low approval of President Macron; and in the UK, Brexit confusion and the current political impasse appears to be sufficient.

Strategic Outliers: Operations Against Smaller Member States

With this new context of social tension, it makes sense why some countries with lower MEP allocations are suffering disproportionately high misinformation penetration on social media.

Holland stands out as a prime example, but also fertile terrain given political tension around the rise of the Party of Freedom. At the time of writing, the recent attack on a tram in Utrecht was not included in this analysis, but will likely provide ample fodder.



Seizing & Shaping Real World Events

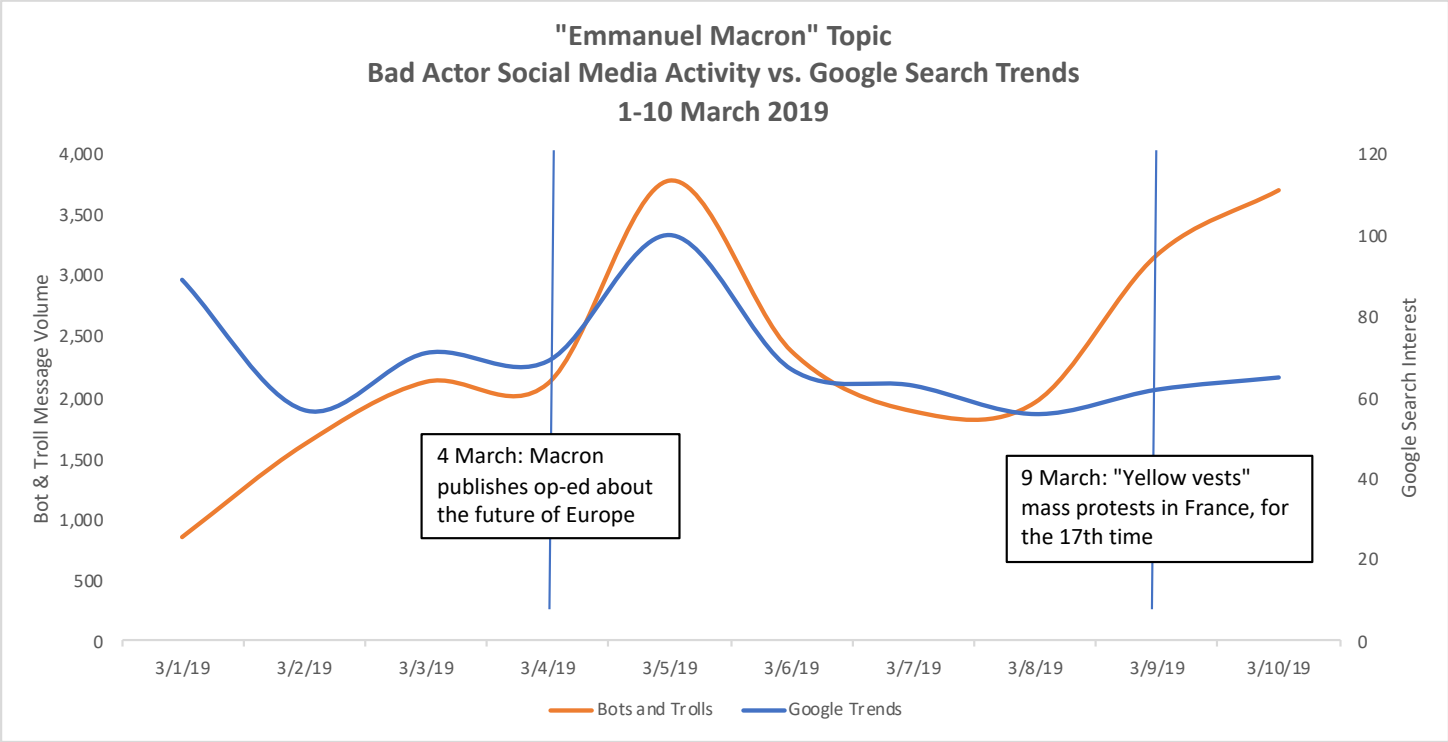
As we have noted in earlier reports, bad actor activity is always-on, but not uniformly. At times, some bot “battalions” and trolls appear to go dormant, tweeting only enough to maintain the charade of an active user. Similarly, we have observed how groups of bots were “activated” at appropriate moments, based on intent.

Misinformation operators mobilize quickly around real-world events either to 1) shape public perception, or 2) project narratives into the zeitgeist that suit their interest in exacerbating social division.

In the March snapshot, we can see both of these scenarios play out against events in France. On 4 March 2019, French President Emmanuel Macron published an article about what he perceived to be the future of Europe. Bad actor activity increased 79% on 5 March, from the previous day, mostly to promote or share content attempting to discredit his ideas. This was a case of trying to shape public perception.

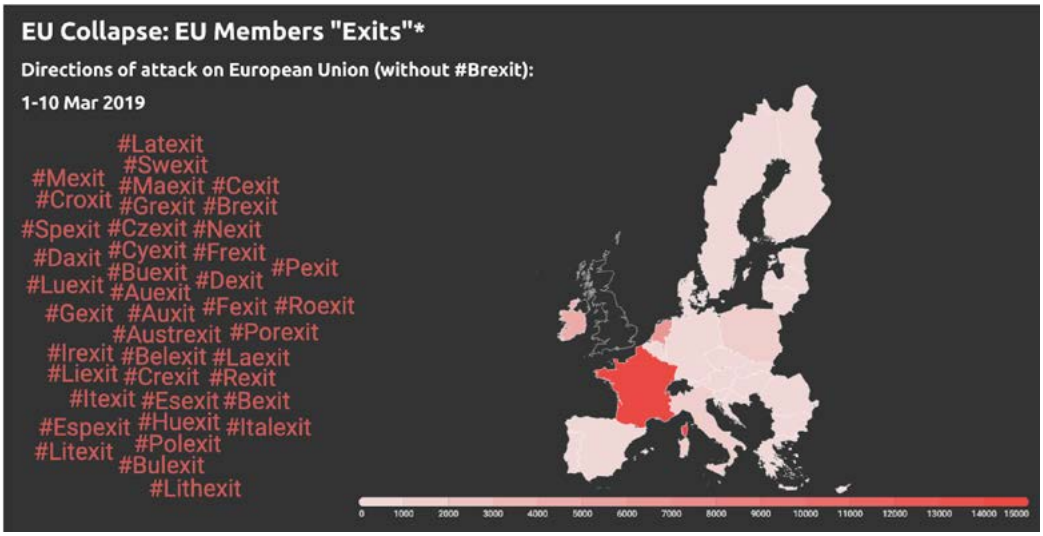
In the period of this snapshot, we also observed another “yellow vests” protest on 9 March 2019, that received only modest coverage in the global press. In this case, a resulting 62% increase in bad actor activity was trying to portray the lack of coverage as part of a larger conspiracy to silence “yellow vest” protesters.

The relationship between misinformation operations and these examples is more apparent when bad actor activity is graphed alongside Google search trend data.



Pushing for the #Exits

Throughout this total time period bots and trolls have also been pushing and amplifying content within other EU member states encouraging other exits (e.g., #Dexit for “Deutschland Exit”, #Frexit, etc.). This content has been always on since Brexit, but in this March snapshot, it’s clear how the surge in activity above corresponds with that messaging volume when we map the #Exit content geographically.

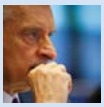


France led the way with 15,224 mentions, while the Netherlands, noted as an apparent outlier, follows closely with 6,347 during this same time period.

Risks for EU Personnel & Digital Infrastructure

Ensuring election integrity also means ensuring the integrity of EU personnel. Recent state-sponsored hacking operations have infiltrated organizations via employees' social media accounts.

We ran a sample of EU officials' Twitter accounts through our Threat Chain analytics engine to get a sense of whether bad actors were not just disseminating content but perhaps targeting EU government officials. We chose Twitter because these accounts represent official, public-facing communications channels, which represent high value targets for account compromise. This sample of users was sourced across the bureaucratic hierarchy. It should be noted that these individuals and their titles were located with cursory searches on LinkedIn, highlighting the ease with which bad actors could also source potential targets.

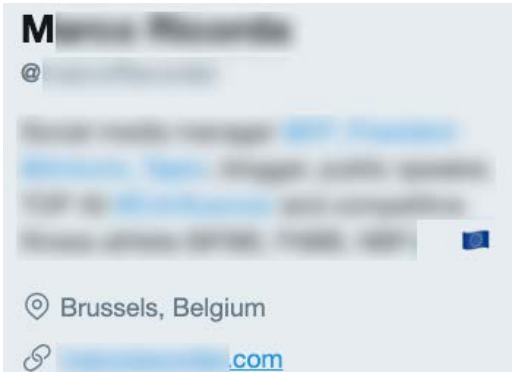
	Name	Job	Bad Actor Activity
	Jean Claude Juncker	President of the European Commission	12%
	Sir Julian King	European Commissioner for the Security Union	13%
	Antonio Tajani	President of the European Parliament	12%
	Jeppe Kofod	Danish MEP	16%
	Jerzy Buzek	Polish MEP, former Polish PM	20%
	Despina Manousos	Policy Adviser, European Parliament	4%

Looking at this chart, for example, 12% of the accounts following Jean-Claude Juncker's official Twitter profile register as bad actors. To be clear, that's 12 out of every 100 followers. Of greater concern is the level of malicious content posted by these followers, in conjunction with misinformation content. Alarming though that may be, perhaps the danger of account compromise is lessened by the layers of approval needed through his media team. However, individual MEPs show a high-level of danger, as do frontline employees, such as those on official media teams.

Why This Matters

In the case of one employee not featured in the chart above, they are on the media team for a high-ranking public EU official. This employee was easily located on LinkedIn using a cursory search.

On this employee's Twitter profile, they provide a link to personal website and descriptions of their interests and pastimes:

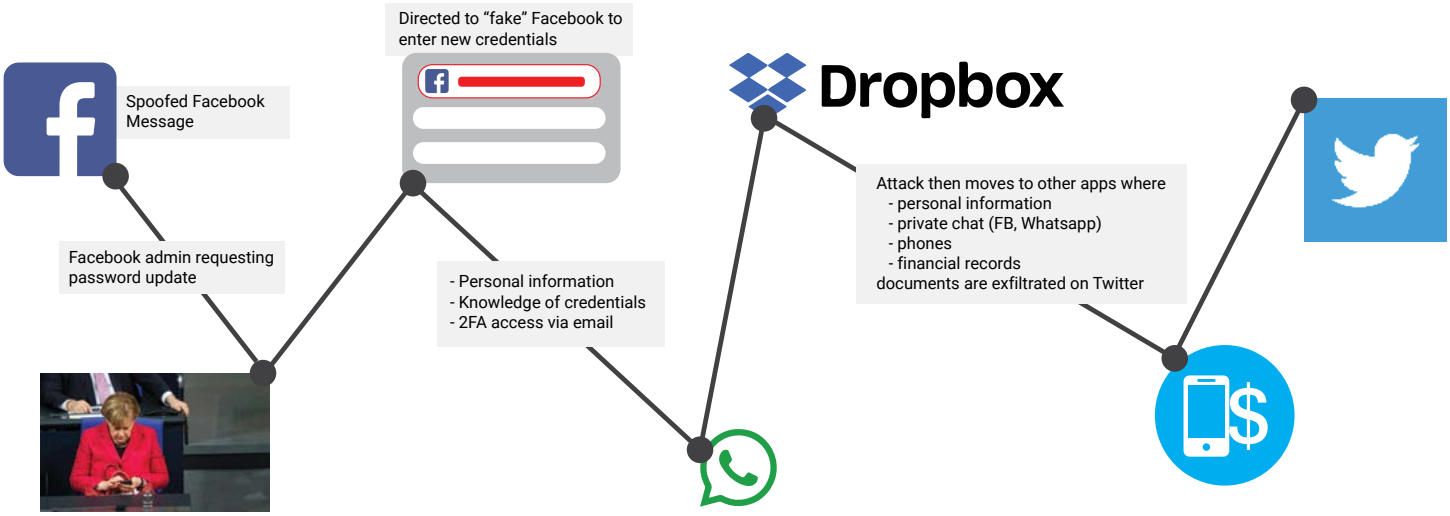


With this username and interest information, it took our investigators fewer than two minutes to locate their personal account on Instagram.

Therefore, for bad actors Juncker and other high-ranking officials are not the targets, this lower-level employee is.

Bad actors who are following this individual, (16% of their Twitter followers register as bad actors) with whom they may have already exchanged messages or likes, could easily launch targeted social engineering attacks against this employee. With one account, they could hop to another until they compromise an official EU channel. From there, it could be simple defacement, or they could send a Direct Message posing as another high ranking official to compromise more accounts, and so on.

This is not a hypothetical scenario. This method of attack is precisely how 20-year old hacker, 0rbit, was able to compromise the accounts of hundreds of German politicians and public figures in December 2018. Posing as Facebook, he sent targets a request for a password update, directing victims to a phishing site disguised as an official Facebook webpage. With stolen credentials tied to emails, 0rbit was able to bypass two-factor authentication protocols. From there, 0rbit replicated the strategy to account-hop to other, more sensitive, channels like WhatsApp, DropBox, and even into victims' banking apps. Collecting sensitive information for months, 0rbit completed the circle as it were, by exfiltrating the data via Twitter.



Employees' exposure to bad actors via official and personal social media accounts dramatically expands the attack surface that can be exploited to influence election operations.

Conclusion

With this report, we set out to cast light on the framework that Russia is employing for social media misinformation operations in the European Union. We hope that our findings will contribute to helping build information and cyber defenses that are commensurate with the level of offense we have observed.

To ensure the integrity of the electoral process, the European Union must look to harden its defenses on two fronts: disinformation and social-engineering cyber attacks. For the first, it is insufficient to secure voting machines if the war is being waged against voters' minds and perceptions. Rigorous fact checking, publication of fact-checking findings, and education around how to spot misleading stories are required. We are confident that Europe is up to the task of marshalling such resources to combat misinformation. Sadly, history shows us that similar measures were required to counteract wartime propaganda during the last century.

Secondly, the European Union must look to secure both high-profile individuals (traditionally referred to as VIPs) and its most vulnerable people (what we call MVPs). Personnel are the first line of defense against social media cyber attacks. Recent events have shown us just how easy it is to spear-phish people on social media, and use the intrusion to gain access to more sensitive data and systems. This front is new and unique to the 21st century, and will require technology that can confront the scale of digital communication channels.

SafeGuard Cyber

Americas
410A East Main St.
Charlottesville, VA 22902
USA

+1 (800) 974-3515

sales@safeguardcyber.com

Asia-Pacific
PO Box 523
Leichhardt NSW 2040
Australia

+61 (437) 276-739

APACsales@safeguardcyber.com