978-9934-564-91-8

DASHBOARD
My profile
Users 242 -4.5% last week
New users 55 +1.8% last week
Page views 908 -2.9% last week
Avg. time 5:36 +9.5% last week
Bounce rate 34% -6.2% last week
Conversion 12% +3.7% last week
Visitors countries
Traffic origin
Target completion
72%
45%
Sales
$23K
Visitors devices
Downtime
+4.1%
SEND MONTHLY REPORT BEFORE 02/03
CALL ROBERT
MEETING MONDAY 15:20
Answer Anna

SOCIAL MEDIA MONITORING: A PRIMER

Methods, tools, and applications for monitoring the social media space

Published by the
NATO Strategic Communications
Centre of Excellence

ISBN: 978-9934-564-91-8
Author: Henrik Twetman, Marina Paramonova, Monika Hanley
Contributors: Rolf Fredheim, Kristina van Sant, Giorgio Bertolini, Alvin Lim, Sebastian Bay
Copy Editor: Anna Reynolds
Design: Linda Curika

Riga, December 2020
NATO STRATCOM COE
11b Kalnciema Iela
Riga LV1048, Latvia
www.stratcomcoe.org
Facebook/stratcomcoe
Twitter: @stratcomcoe

This publication does not represent the opinions or policies of NATO or NATO StratCom COE.

PAGE LEFT BLANK INTENTIONALLY

PAGE LEFT BLANK INTENTIONALLY

Contents

”

Developments in technology, communication and demographics all shape and alter the character of conflict. Information flow is now so prevalent, potent and unavoidable that it forms as much a part of the operations environment as the terrain or weather.”

NATO AJP 3-10.11[1]

INTRODUCTION

For the last twenty years, the phenomenon of social media has cemented itself as the new nexus of social interaction. With billions of users across the world, hundreds of online platforms, and a myriad of digital technologies at its backbone, social media is fundamentally reshaping our understanding of the global information environment. Today, social media is essential infrastructure for personal conversation, public debate, and commercial communication.

While social media provides many opportunities for unprecedented information sharing, the rapid adoption of limitless communication technologies with instant amplification and global reach has also created significant vulnerabilities. Social media have, in many cases, become a conduit for unsubstantiated information, such as rumours, hoaxes, and conspiracy theories.[2] Even more worrisome from a national security perspective, hostile actors are deliberately exploiting social media to spread disinformation and to conduct information influence activities with the intent of deceiving and misleading audiences to achieve their strategic aims.[3] Contemporary conflicts, especially those that fall within the hybrid spectrum,[4] increasingly play out over social media,[5] and actors such as Russia, China, Iran, and Saudi Arabia have recently expanded their efforts to manipulate the social media space.[6]

"**Social media** is understood as the different forms of online communication used by people to create networks, communities, and collectives to share information, ideas, messages, and other content."

Erragcha & Babay (2020) Social Media, Marketing Practices, and Consumer Behavior

Malicious use of social media poses a clear security challenge. The opportunities provided by new digital technologies are exploited to undermine trust in democratic institutions and legitimate news sources, to distort public discourse and opinion formation, and to influence elections and short-circuit decision-making processes.[7]

For this reason, it is important for a wide array of stakeholders to understand what is happening on social media. Effectively listening to conversations online and monitoring and analysing social media content is crucial for both public and private sector actors,[8] as well as for military organisations.[9]

This, however, is easier said than done. The complex social and technical infrastructure of the online environment makes both seeing the big picture and identifying specific pieces of information challenging. The speed at which information flows between social media users, and the ever-changing types of data generated further complicate the picture.[10] The deceptive nature of disinformation[11] and information influence activities also make detection and attribution difficult.[12] It has quickly become clear that attaining perfect situational awareness of the online information environment is a tall order.[13]

Still, while much of today's information environment is essentially characterised by perpetual chaos, it is possible to study and understand it, albeit momentarily. Even snapshots of a bigger picture are critical for operating successfully in this space. The importance of understanding what is happening on social media has prompted the development of a wide range of tools to monitor, measure, and analyse metrics and content. Oftentimes these tools are developed with either a commercial objective—monitoring brand engagement or customer discourse—or with a scholarly mindset—to understand wider patterns and trends. However, these tools can also be leveraged to gain insights regarding disinformation and other security concerns.

A primer on social media monitoring

This primer sets out to achieve three objectives. First, it seeks to provide an overarching perspective on social media monitoring and the complexities of studying the online information environment. Second, it provides insights into how monitoring can be structured and conducted to achieve desired results.

“NATO views **disinformation** as the deliberate creation and dissemination of false and/or manipulated information with the intent to deceive and/or mislead. Disinformation seeks to deepen divisions within and between Allied nations, and to undermine people’s confidence in elected governments.”

NATO (2020) NATO’s approach to countering disinformation: A focus on COVID-19

Information influence activities

“Information influence activities are activities conducted by foreign powers to influence the perceptions, behaviour, and decisions of target groups to the benefit of foreign powers.”

Pamment et al. (2018) Countering Information Influence Activities: The Start of the Art, MSB

Finally, it provides an overview and review of some popular tools for social media monitoring from a disinformation perspective, including tools for data collection, network analysis, and data visualisation.

The primer serves as a point of departure for developing a social media monitoring capacity, with specific reference to security issues such as information influence activities and disinformation. We hope that our overview of methods and tools will provide an entry point for beginners and a new perspective for intermediate and advanced users, be they open-source operators, public sector communicators, journalists, or interested members of the public.

About this report

This report provides an overview of social media monitoring and the complexities of studying the online information environment. It also reviews a variety of tools and services for social media monitoring, particularly focusing on disinformation and information influence.

Research method

This primer builds on consultations with subject-matter experts who routinely work with social media monitoring. We asked developers, intelligence operatives, open-source practitioners, PR experts, journalists, and others, which methods and tools they preferred. We then cross-referenced the information they provided with our own research and in-house testing. The products and services mentioned in this primer should not be viewed as an endorsement or recommendation of the products or services described.

SOCIAL MEDIA

1. SOCIAL MEDIA

Social media is a commonly used term; however, it has no generally accepted definition. This sometimes adds to difficulty in understanding the challenges and opportunities associated with this new form of communication.[14] Simple definitions, such as 'any form of technology aimed at easy intra-personal communication', are too broad, while more specific and technical definitions, such as 'web 2.0 internet-based applications where user-generated content provides the basis for interaction between individuals and groups', are too exclusive.[15]

This chapter attempts to give an overview of the social media space, and to provide nuance and context to the concept of social media monitoring. Understanding what social media is, how it works, and how audiences use it, is a first necessary step for designing an appropriate monitoring strategy.

For the purpose of this report, we understand social media as 'the different forms of online communication used by people to create networks, communities, and collectives to share information, ideas, messages, and other content'.[16] This definition provides a good starting point for delineating the concept, as it highlights some prominent features:

- Social media is online, or internet-based

- Social media enables communities and networks to communicate, discuss, and interact

- Social media revolves around the creation and sharing of information

Social media is thereby different from other types of media (such as paper-based media or electronic media) as well as other types of online platforms. On social media, users are active participants not passive consumers.[17] Social media is characterised by its reach (since the internet is global), its interactivity and immediacy (users within networks communicate directly and almost instantly with each other), and its abundance (anyone and everyone can produce and share information).[18] These characteristics lead to several new dynamics in terms of how communication functions on social media platforms. Shao et al. highlight some of these dynamics, such as homophily, polarisation, and social bubbles.[19]

An abundance of platforms provide social media services. We compiled a list of a variety of platforms that could fall under the umbrella term of social media to illustrate the breadth of these services:

Social Media Platform	Launch year	Headquarters' location	Focus/ purpose	Rank in global Internet engagement[20]	Daily time on the website
Discord	2015	San Francisco, US	Connecting communities, predominantly gamers, but other types as well	121	6:30
Facebook	2004	Menlo Park, US	Connecting friends, family, colleagues and others; sharing pictures, videos, articles; joining groups, etc.	4	18:27
Flickr	2004	San Francisco, US	Photo sharing and management	720	4:30
Gab	2017	Clarks Summit, US	Enabling individual expression	10,729	2:53
Instagram	2010	Menlo Park, US	Sharing photos and videos	30	8:21
LinkedIn	2003	Sunnyvale, US	Networking for professionals	61	10:25
Meetup	2002	New York, US	Finding and building local communities, meeting new people	1,362	3:59
Nextdoor	2011	San Francisco, US	Connecting with neighbours, exchanging information, goods, services	2,535	3:44
Odnoklassniki (OK.ru)	2006	Moscow, Russia	Connecting with classmates and old friends	66	4:17
Pinterest	2010	San Francisco, US	Collecting and sharing images	151	5:53
Quora	2010	Mountain View, US	Asking and answering questions	241	3:55
QQ	1999	Shenzhen, China	Instant messaging; online social games, microblogging, etc.	5	3:46
Reddit	2005	San Francisco, US	Sharing news, ideas, and content in a wide variety of 'subreddits'	19	5:52
ReverbNation	2006	Morrisville, US	Networking for musicians and producers; discovering new music for fans	8,613	4:02
Sina Weibo	2009	Beijing, China	Microblogging; somewhat similar to Twitter and Instagram	16	3:09
Skyrock	2002	Paris, France	Creating blogs; exchanging messages, photos, videos	6,985	2:52
Snapchat	2011	Santa Monica, US	Connecting with friends, sharing live stories and messages that are only available for a short time	3,115	6:14

Social Media Platform	Launch year	Headquarters' location	Focus/ purpose	Rank in global Internet engagement[20]	Daily time on the website
SoundCloud	2008	Berlin, Germany	Sharing music and distributing audio files	104	3:35
Taringa!	2004	Buenos Aires, Argentina	Creating and sharing content; somewhat similar to Reddit	1,332	2:24
Telegram	2013	London, UK	Instant messaging, but also creating communities, sharing content via channels etc.	178	7:05
TikTok	2016	Los Angeles, US	Sharing short videos	272	4:21
Tumblr	2007	New York, US	Microblogging	118	4:21
Twitter	2006	San Francisco, US	Microblogging	48	12:45
VK	2006	Saint Petersburg, Russia	Connecting with friends, family, colleagues and others; sharing pictures, videos, articles; joining communities etc.	23	8:15
We Heart It	2008	San Francisco, US	Sharing images; similar to Pintrest	2,528	4:48
XING	2003	Hamburg, Germany	Networking for professionals	1,845	3:55
YouTube	2005	San Bruno, US	Sharing videos	2	13:44
YY	2008	Guangzhou, China	Streaming and sharing videos; features a virtual currency which can be converted to real currency	70	2:49

As the table illustrates, 'social media' can refer to a variety of platforms providing different services. The definition suggested above also includes social messaging platforms—direct messaging services that provide distinct social functions. These platforms are increasingly popular as more and more users favour direct communication over more public platforms such as Facebook or Twitter. Today, social messaging platforms account for a combined 4.1 billion users; social messaging is the most frequent activity a person carries out online.[21]

Social Messaging Platform	Launch year	Headquarters' location	Distinctive characteristic	Estimated number of active monthly users[22]
BAND	2012	Seongnam, South Korea	Chat app for groups with features such as a community board, shared calendar, polls, etc.	Over 2.5 million
Dust	2014	Los Angeles, US	Privacy and security focused app that positions itself as an encrypted messenger, a stealth search engine (no cookies, no tracking), and a watchdog (see if your data has been compromised before/get alerts if it becomes compromised)	N/A
Facebook Messenger	2011	Menlo Park, US	Seamless integration with Facebook accounts	1.3 billion
Line	2011	Tokyo, Japan	Hidden chats, gamification; created by the Japanese arm of Naver Corporation, which also created BAND	217 million
Kik	2010	Santa Monica, US	Anonymous messaging; use and build bots for a variety of purposes—to chat, do quizzes, get news and advice, etc.	15 million
KakaoTalk	2010	Jeju-si, South Korea	Voice filters, mobile games	50 million
QQ	1999	Shenzhen, China	Online social games, microblogging, shopping, music, etc.; developed by Tencent	731 million
Signal	2014	Mountain View, US	Privacy and security focused; non-profit, not tied to any major tech companies; open-source	N/A but the app has been downloaded over 10 million times
Slack	2013	San Francisco, US	Focused on business communication and collaboration.	N/A but over 12 million daily active users
Telegram	2013	London, UK	Direct chats, broadcasting public messages to large audiences, bots	400 million
Threema	2012	Pfäffikon, Switzerland	Can be used anonymously, no need for a phone number or any other personal information; generates little user data; a paid app	N/A but a total of 5 million users in 2018
Viber	2010	Limassol, Cyprus	VoIP and messaging app	260 million
WeChat	2011	Shenzhen, China	News, games, mobile payment features, local services (various third-party services), etc.	1.15 billion (combined with Weixin)
WhatsApp	2009	Menlo Park, US	Encrypted messaging app; owned by Facebook.	2 billion
Wickr	2012	San Francisco, US	No phone number or email address needed to register; auto-destruct feature for messages, with expiration and 'burn-on-read' timers available	N/A

TYPES OF SOCIAL MEDIA

There are a several different types of social media, each with its own internal logic, which must be understood for monitoring to be meaningful. Different users engage with different types of platforms, and different platforms contain different content. To highlight some of the differences, we have provided a basic categorisation of platforms and services, all of which are referred to as social media. [23]

Types of social media:

- Social networks
- Media-sharing networks
- Social messaging platforms
- Blogging networks
- Discussion boards
- Review networks

Social Networks

Social networks are what we usually think about when we say 'social media'. These are online platforms designed to connect people and organisations with each other, creating networks. These platforms often provide additional features but are primarily designed around connecting people and enabling the sharing of information between them.

Disinformation and information influence actors frequently take advantage of the social nature of these platforms, for example by exploiting the phenomenon known as 'social proof'—the impression of a social setting makes users susceptible to disinformation because others seem to believe it.[24] This is achieved, for example, by setting up fake profiles to engage others in discussion, by creating false groups, or by sharing misleading content or links to disinformation.

Heart of Texas Facebook Group[25]

There are countless examples of disinformation and information influence activities on social networks, but one of the more well-known cases illustrates the structure of such activities. In the leadup to the US presidential election of 2016 a series of fake accounts and fake groups were set up on Facebook and other platforms to polarise domestic debate. One of these groups was 'Heart of Texas'—a fake US secession group run by the Russian Internet Research Agency, which at its peak had more followers than the official Texas Democratic and Republican party pages combined.[26]

Media-sharing networks

Media-sharing networks are designed specifically around sharing media, such as images (Instagram) or videos (YouTube), but they have a distinct social element to them as well.

As these networks are audio-visual in nature, they are excellent for sharing disinformation in the form of, for example, memes or misleading video-clips, but also for metric and comment-based manipulation (views, likes, and shares) to inflate popularity and misrepresent affiliation.[27]

Hong Kong protest disinformation on YouTube

During the peak of the Hong Kong protests in 2019 the video-sharing platform YouTube identified and disabled 210 channels originating in China for spreading disinformation related to the protests. According to the YouTube, the channels had acted in a coordinated manner that was consistent with disinformation campaigns observed on other platforms such as Facebook and Twitter.[28] Curiously, after this disinformation campaign had been shut down by the major social media platforms, Quartz reported that disinformation resurfaced on the pornographic network Pornhub.[29] This illustrates how platforms not normally considered in this context can be exploited for influence purposes.

Social messaging platforms

While many other platforms are open, social messaging platforms are defined by their closed nature. For this reason, they are sometimes not considered to be social media platforms. We see them as social in nature due to distinct social features of apps such as WhatsApp and Telegram, that allow for large discussion groups and easy sharing of content.[30]

Because these platforms are closed, and often encrypted, monitoring is difficult. For this reason, they are frequently exploited by disinformation operatives to circulate misleading messages or questionable links to large audiences.

DAESH propaganda on Telegram[31]

DAESH has been known to use the encrypted social messaging platform Telegram to spread propaganda and disinformation. The platform has been convenient for this purpose as it allows for the creation of channels where selected users can interact with each other. Propaganda distributed on Telegram typically transitions to wider audiences on other platforms, such as Twitter, via crowdsourcing where individual users independently repost the content elsewhere.

Blogging networks

Blogging networks, or blogs, are online spaces where people can express and publish their ideas and thoughts. Blogs allow for more complex and lengthy information compared to most other social media platforms. Blogs have a strong social element despite not always having an infrastructure for engagement beyond comments, as blog posts are frequently shared and disseminated across other platforms. Blogs can either be self-hosted or hosted on a shared platform, such as Tumblr.

Blogs are useful for framing narratives[32] and can provide a force-multiplier effect to campaigns run on other platforms, such as Facebook or Twitter, or be used to legitimise a Potemkin village of evidence.[33] Blogs have previously been observed in wider information influence activities and disinformation campaigns. [34]

5G disinformation on blogs[35]

The debate around 5G technologies is ridden with disinformation, oftentimes reinforced by state actors who seek to influence public discourse negatively. The Global Disinformation Index identified a strategy for anti-5G disinformation campaigns. It begins with the establishment of an adversarial narrative, which is supported by web artifacts, such as blogs, that back up disinformation claims with false or misleading evidence. Such 'narrative incubation' precedes the establishment of a disinformation narrative disseminated through social networks and provides a point of departure for the disinformation campaign.[36]

Discussion boards

Discussion boards, or forums, are likely the oldest type of social media, predating social networking and other messaging platforms by decades. Discussion boards typically engage niche audiences and are focused on specific topics and themes. The larger discussion boards, such as Reddit and Quora, are designed to support huge communities and a wide variety of topics, while smaller boards focus on specific audiences.

The niche nature of discussion boards, where people join to discuss their common interests, makes them susceptible to disinformation and information influence activities, as they provide very clear target audiences with clear interests and concerns. Discussion boards often provide anonymity for their users, lowering the threshold for manipulation.

Election document leak on Reddit[37]

In late 2019, a user on the discussion board Reddit posted leaked UK government documents to influence UK domestic politics. A digital forensic investigation by Reddit revealed that the leak appeared to have originated in Russia and was part of a wider influence operation involving a series of connected accounts that reposted the document across multiple forums in an attempt to manipulate the platform's upvoting system to gain popularity.[38] When discovered, Reddit banned one subreddit and 61 accounts under their policies against voter manipulation and misuse of the platform.[39]

Review networks

Like social messaging platforms, review networks are not always included in the social media family, but they too have distinct social elements that can be exploited. These networks are used to provide reviews and assessments of brands, products, experiences, services, or anything else. Consumer reviews are valuable and impact our perceptions of the product or service under review.

Such networks have not played a major part in disinformation campaigns or information influence activities to date, but they have proven useful conduits for influencing both the perceptions and behaviours of target audiences in other contexts. The coordinated manipulation of reviews through, for example, purchasing fake reviews from a social media Manipulation Service Provider,[40] has been repeatedly observed on larger review networks.[41]

Manipulation of TripAdvisor by a Vice reporter[42]

While not strictly speaking a case of disinformation, Vice journalist Oobah Butler's campaign to create a top-rated restaurant in London out of his backyard shed illustrates the potential for manipulation via review networks. Using a burner phone, a series of fake reviews, and a smart communication strategy, Butler managed to trick the TripAdvisor platform into ranking his shed the premier restaurant in London and had a stream of guests calling to make reservations.

Further reading

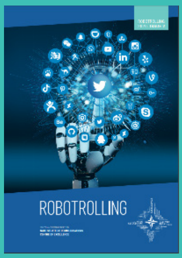
Since its foundation, NATO StratCom COE has studied social media manipulation as an important part of the influence campaigns malicious state and non-state actors direct against the Alliance and its partners. Here are some reports that provide a deeper understanding of how the online environment is exploited and manipulated.

Falling Behind: How social media companies are failing to combat inauthentic behaviour online

Social media platforms are ridden with inauthentic users and accounts. Oftentimes these are created by Manipulation Service Providers (MSPs) who have commercialised the manipulation of social media by setting up complex infrastructures of exploitation. Many of these providers are based in Russia and sell social media engagement in the form of comments, clicks, likes, and shares. Researchers spent €300 to test how these services work and to see how social media platforms are handling inauthentic engagement.

Manipulation Ecosystem of Social Messaging Platforms

Messaging services are also vulnerable to manipulation. In an effort to understand how this works, we examined how WhatsApp and Telegram, two of the most popular messaging services in the world, can be manipulated. The authors conducted an online investigation into the quality and range of these services and spoke to sellers and freelancers to understand what a malicious actor could easily purchase online. They looked at the cost, methods, and quality of manipulation services.



Robotrolling

Automated accounts, also known as bots, often drive the spread of disinformation on social media. The quarterly report monitors automated activity on Twitter and VK to discern trends and patterns. The reports are based on data collected by a bespoke AI system that scrapes these platforms for information. Since 2017, *Robotrolling* has reported on bot use related to current issues in the Baltic states and other priority areas for NATO.

Camouflage for the Digital Domain: A force protection framework for armed forces

Social media is but one part of the digital domain. This report, produced together with NATO Cooperative Cyber Defence Centre of Excellence (CCDCOE) in Tallinn, examines the broader risks of the digital domain, putting disinformation and information influence activities into a broader military context and discusses a framework for mitigating digital risks.

USER VARIATIONS AND PATTERNS

It is not just platforms that vary; who uses them and how they are used do also. From analysing user statistics, it is clear that audiences choose certain platforms for certain purposes, and that factors such as gender, age, region, political opinion, socioeconomic status, and device use, all impact the choice of platform or service.

Not only is there variation among patterns of use, but the online environment is constantly shifting. A platform or service can gain immense popularity in no time and disappear just as quickly. The early social networking site MySpace is a good example. It was the most visited website in the world in 2006 and was traded at a staggering $580 million the year before. Just two years later the platform was surpassed by Facebook and was later sold for only $35 million.[43] Today, Alexa gives it a rank of 2,417.[44]

Understanding these varying patterns of social media usage is important for monitoring, not least to recognise selection bias when studying information from a specific source (more on biases in Chapter 4). Studying the spread of disinformation on Twitter, for example, is highly relevant in a country such as the US where Twitter is a popular platform for politicians, journalists, and other elites, whereas it may be less revelatory in a country where other social media platforms are more widely used.[45]

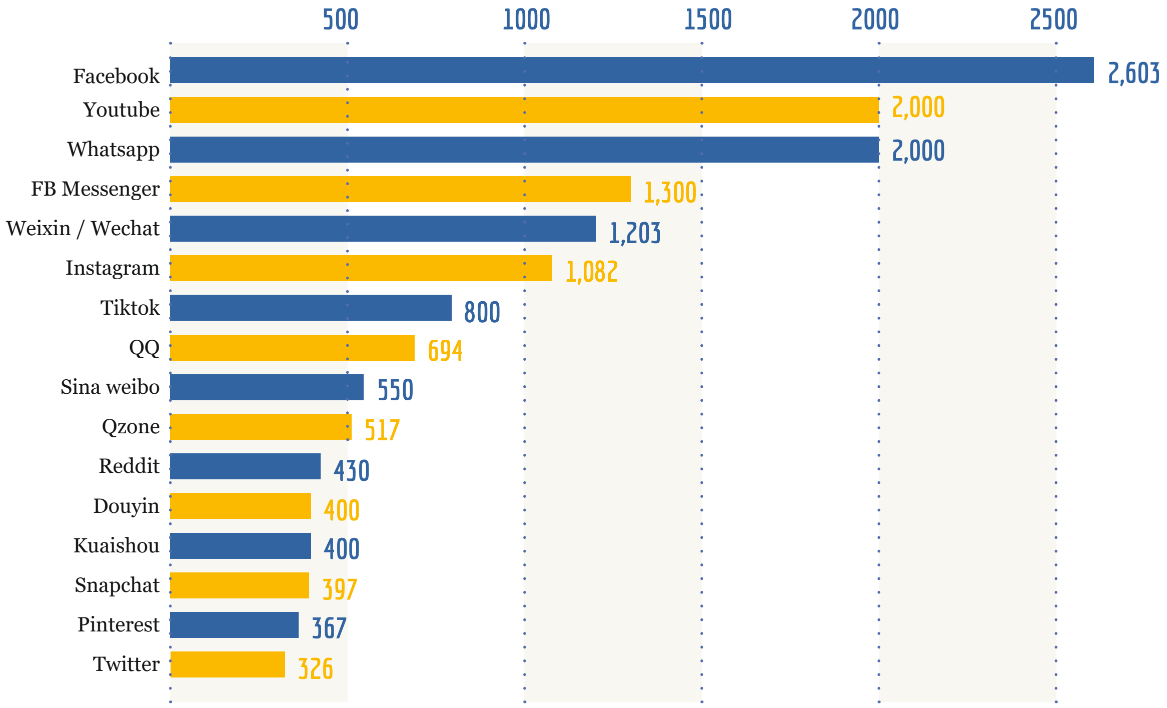


Figure 1 The world most-used social media platforms - in millions (Source: Hootsuite, 2020).

Regional Variations

Depending on where in the world you focus, the social media landscape will look different, even if the giants dominate globally. Facebook and Instagram are the two most popular social media platforms in most countries, with the notable exceptions of Russia, China, and parts of central Asia and the Middle East.[46] But greater divergence emerges if we look beyond these two.

Not only do preferred platforms vary between regions and countries, internet access and social media penetration also differ. A report from Statista (January 2020) shows, for example, the UAE, Taiwan, and South Korea as having the highest social media penetration in the world—all above 87%—and Nigeria, Kenya, and Ghana having the lowest—all below 20%.[47] In developing countries where internet infrastructure has limited capacity, data-light apps such as Facebook Light and Twitter are often favoured over heavier apps, such as Instagram.

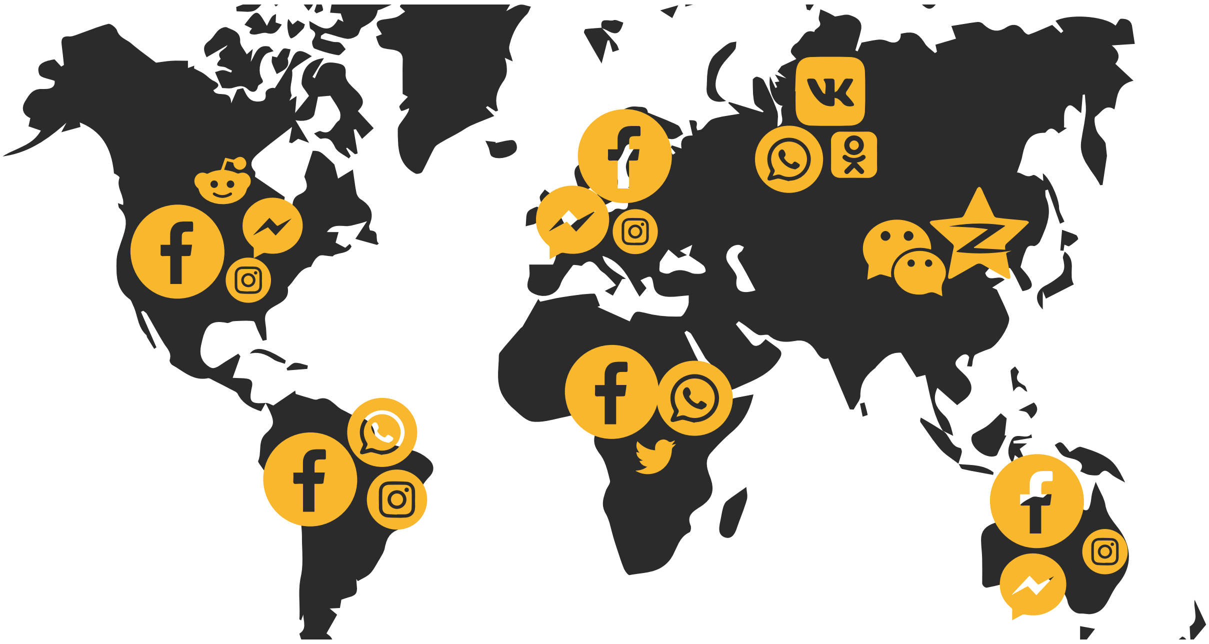


Figure 2 The popularity of different social media platforms differ by region (Aggregation based on World Map of Social Networks)[48]

Demographic Variations

Regional statistics give a broad indication of social media consumption patterns, but different population groups within regions use social media differently. Studies indicate that demographic, structural, political, and socio-economic differences are mirrored in the use of social media platforms and other internet services.[49] People also use different platforms for different things, adding further to this complexity. Differences in social media use often reflect differences in the offline space.[50]

Among demographic variables, age strongly impacts social media use. Figure 3 illustrates not only that different generations use social media services to different extents, but also that they prefer different platforms. Younger generations use media-cantered platforms such as Instagram, TikTok, and Snapchat, whereas older generations are more active on social networking sites such as Facebook and Twitter. However, this picture is changing rapidly—the World Economic Forum notes that the so-called boomer generation is continuously increasing its activity on social media and is starting to expand into different types of social media such as messaging platforms and media-sharing networks.[51]

For younger generations, social media is also, to some extent, competing with search engines as a means to research, for example, products and brands.[52] Older generations seem to focus more on the communicative aspects of the platforms. Younger people tend to be early adopters of new platforms and trends, whereas older people move more slowly between services.[53]

Other demographic variables such as gender and race also matter. While both men and women use social media at similar rates, they favour different platforms. Women, for example, are nearly three times as likely as men to use Pinterest,[54] whereas 62% of YouTube users are male.[55] Research done by the Pew Research Center indicates that, at least in the US, different ethnic communities favour different platforms. WhatsApp, for example, is used by 42% of the Hispanic population but only 13% of whites.[56]

Similar variations are observable for different socio-economic groups, with those in higher-education and higher-income groups using different platforms and services compared to lower-education and lower-income groups.[57] Similarly, urban populations are more active on social media than rural populations.

Variation in use by demographic factors tells only part of the story. Variations also occur at the level of personal relationships, political affiliation, and group belonging. The social media platform Gab illustrated this dynamic, as its focus on preservation of freedom of speech has attracted niche communities with particular political alignments that perceive themselves as disenfranchised from other platforms.[58]

Current Trends

Taken together, these variations paint a complex picture of the social media landscape where various groups use different platforms for assorted ends depending on myriad variables. The landscape is also constantly shifting, and awareness of how it does so is important for monitoring.

Accurately capturing the constant transformation of this complex space is impossible, but some general trends can be discerned. One such trend is the move towards increased privacy and integrity on social media.

Future social media landscape

"Today, we already see that private messaging, ephemeral stories, and small groups are by far the fastest growing areas of online communication."

Mark Zuckerburg, 2019[59]

This implies a shift from large public groups with open content to private forums, encrypted services, disappearing (or ephemeral) content, and smaller groups.[60] This can already be seen, for example, by the increasing popularity of encrypted messaging apps such as WhatsApp, and the rise of apps and functions with ephemeral content such as Snapchat and the 'Stories' feature on Instagram and Facebook. From a monitoring perspective this poses challenges, as collection and analysis of user histories will be much more difficult.

Similarly, niche apps designed for specific communities with particular interests or views are becoming more popular and will likely play a more prominent role in the future.[61] This will lead to even greater diversification of social media, which could lead to the emergence of even more filter bubbles and echo chambers.

Disinformation on Social Media

Due to its unique dynamics, social media has become a conduit for disinformation and hostile influence operations. Social media enables disinformation operatives to quickly reach large numbers of people, hide their identities, use technical manipulation to automate their efforts, and target their communications to specific audiences. The social dimension of social media also benefits actors spreading disinformation as ordinary users who engage with it act as force multipliers—perpetuating, spreading, and validating the disinformation.[62]

Disinformation on social media comes in many shapes and forms. It can manifest as a piece of fake news that is shared, a deliberately false post or a comment, a network of automated accounts pushing a specific narrative, a misleading meme, an anonymous user engaged in trolling, or a targeted advertisement with misleading content. Oftentimes, disinformation is not verifiably false but is misleading, nonetheless.

Actors who produce and spread disinformation on social media use established infrastructures[63] and exploit existing issues and polarised audiences; this gives them an advantage.[64] These actors tend to be agile and flexible.[65] They use multiple techniques, engage in trial-and-error, make use of multiple platforms, and are not bound by any fixed modus operandi (although some techniques recur frequently).[66]

The scope of disinformation is often unknown, and we often fail to see the big picture until it is too late. Disinformation spread during a specific political event may not be aimed at influencing current developments but might be designed to achieve a longer-term effect.

The UK's guide to tackling disinformation, RESIST, offers a useful taxonomy of that illustrates how such phenomena appear on social media. The five 'FIRST Principles of Disinformation' can be used to delineate your monitoring, defining which indicators you need to look for.

FIRST Principles of Disinformation[67]

Fabrication: Disinformation often manipulates or fabricates content, such as including untrue information, attaching a manipulated picture, or sharing a doctored video.
Identity: Disinformation frequently makes use of false identities and sources, such as an anonymous or fake account.

Rhetoric: Disinformation is not necessarily false but can also make use of malign or false arguments to skew a discussion. Trolling is an example of how harsh rhetoric can be used to this end.
Symbolism: Disinformation leverages and exploits events for their symbolic and communicative value.
Technology: Disinformation exploits a technological advantage, for example, by using automated accounts to amplify their spread.

For more information, see: GCS (2019)
'RESIST: Counter Disinformation Toolkit'[69]

The following publications provide good entry points for understanding how disinformation works on social media

Countering Information Influence Activities: The State of the Art	Swedish Civil Contingencies Agency	Overview of current thinking on how to counteract information influence activities
RESIST Counter Disinformation Toolkit	UK Government Communication Service	Guidance for how to recognise and counter disinformation
How Do You Define a Problem Like Influence?	Carnegie Endowment for International Peace	Research article on the phenomenon of influence operations

SOCIAL MEDIA MONITORING

2. SOCIAL MEDIA MONITORING

Social media monitoring refers to the tracking or observing of online social media and other websites for relevant information. Whether for business or disinformation research, social media monitoring is the process by which an individual or a company can 'listen' to social media conversations pertaining to a topic they are interested in. This can be done manually, but logging into individual social media platforms and reading content you are interested in, but it usually involves some degree of automation or the use of third-party software to speed up the process and enable collection of information on a larger scale. Like search engines, social media monitoring normally involves the use of algorithm-based tools that can go through social media sites and gather data that is then indexed for use.

Monitoring can involve a variety of tasks including mentions of a topic or company, hashtag tracking, finding influencers, and trends regarding the topic you wish to monitor. With the data collected, one can then move to organise, analyse, and look for insights.

This chapter introduces what social media monitoring is and how it works. Coming to terms not only with how interactions on social media can be observed and studied, but also with the limitations of this work, is crucial for setting up a monitoring system for your own needs.

Monitoring or listening?

'Social media monitoring' and 'social media listening' are sometimes used interchangeably. Generally, social media monitoring can be said to be a part of social media listening. Monitoring involves more of the immediate collection of data without depth. For example, if an event takes place and a new hashtag is formed, your monitoring could tell you how it started, how it spread, and how the hashtag is being used. However, it may not give you the deeper analysis that is gained over time using social media listening, defined as 'social media monitoring plus analysis and insights or data visualisation'.[69]

HOW SOCIAL MEDIA MONITORING WORKS

Social media monitoring is an iterative process, as it is part of a cycle of information collection (see Chapter 4). Generally, social media monitoring works by crawling social media platforms to retrieve data, and indexing the results, which can be done in real time or at certain set time intervals. This then allows for establishing an alert system to keep up to date on changes, emerging trends, or developments in the topic or keyword you are following.

Social media monitoring tools generally query social media platforms by means of the platform's API (application programming interface), which allows the tools to interact with the social media platform using a predetermined standardized language. API access enables third-party software to perform a range of activities. This may be as simple as scheduling social media posts via another programme, or as complicated as rapid data access based on specific queries.

Each social media platform restricts the information social media monitoring tools can access, and these restrictions are frequently changed. For example, VKontakte, the Russian social media platform, has a very liberal and open API. On the other hand, Facebook currently allows only public pages to be queried, while personal accounts are inaccessible. Differences in APIs often determine whether your insights can be qualitative or quantitative, depending on level of restriction and access.

Some of the major social media companies offer free use of their API; others require special arrangements. Twitter, for example, have a relatively open API but offers more enhanced access through a private paid agreement. Rate limits are in place to help with scalability and to ensure that the API is not slowed by too many requests at once. Rate limits can be user-based, time-based, or server-rate-limiting. Some platforms have different rate limits depending on type of query. Below is an overview of some of the most popular social media platforms:

- Instagram: (access token, OAuth 2.0) limited demographic information (age, gender, interests) 200 calls per hour, down from 5,000 calls previously.
- Facebook: (access token, OAuth 2.0) 200 calls
- Twitter: (access token) rate limits on a 15-minute basis
- VKontakte: (access token, OAuth 2.0) rate limits of at least 3 calls per second
- Telegram: (access token, OAuth 2.0) 100 requests per second
- Pinterest: (access token, OAuth 2.0) 1,000 calls per endpoint per hour
- YouTube: (API key or OAuth 2.0) 2.0 40,000 calls per day [70]

Social media monitoring can also identify sources or originators of information, influencers, or spreaders of information, and can monitor conversations on certain topics that are ongoing or have an online history. The objective of your monitoring will affect what kind of targets you monitor. Your monitoring system will also likely evolve as you hone your searches and targets and adapt to changes in the social media landscape.

USES OF SOCIAL MEDIA MONITORING

Social media monitoring, in terms of both processes and tools, has primarily been developed and designed with commercial objectives in mind—to monitor brand engagement and consumer relationships. To these ends social media monitoring provides useful assets and insights.

Recently, the relevance of social media monitoring for public sector organisations and security institutions, such as armed forces or intelligence services, has become apparent. Social media monitoring has become a valuable tool in the detection of disinformation and information influence activities.

Tracking disinformation—an example

An illustrative example of how social media monitoring and analysis can be used to detect and track disinformation is provided by the Atlantic Council's Digital Forensic Research Lab (DFR Lab), which used CooRNet—a tool for detecting coordinated behaviour on social media—and Facebook's CrowdTangle—a content discovery and social monitoring platform—to track the spread of the disinformation video Plandemic, which went viral during the Covid-19 pandemic in 2020.[71]

DFRLab obtained a dataset from CrowdTangle indicating which accounts and groups were promoting and disseminating the video by creating a specific search string. They then proceeded to analyse and visualize the data using different tools. Their full report can be read here.

However, as stated throughout, there is no one-size-fits all solution to the challenge of identifying disinformation, and no tool exists yet that will do everything for you. Similarly, social media monitoring will not replace traditional intelligence work, while it may supplement it.

Whether you are a curious individual, a journalist, a government or military official, or a social media marketing manager, you can use social media monitoring tools to keep track of developing stories, topics, conspiracies, or disinformation. Using monitoring tools, one can observe the impact of certain influencers, identify networks, find automated accounts or bots, discover demographic details of groups consuming certain types of information, and receive information about sentiment surrounding a certain topic or conversation. Monitoring is not beneficial only for identifying instances of disinformation but can also provide valuable insights about whether to respond to it, ignore it, or educate the targets of the disinformation so that they can develop resistance to it.

Those who monitor social media in this way are frequently looking not only at the message itself, its source, or the target demographics, but also at the response (positive or negative). Monitoring can provide governments with deeper insight into, for example, the impact a particular piece of disinformation is having on audiences.

Social media monitoring can also provide resources for monitoring the effects of your own communication efforts and inform you about the impact of your competitors in the field. Many companies or organisations use social media listening to track the success or failure of a social media campaign, hashtag use, mentions, and general sentiment surrounding their organisation or brand. Businesses and private individuals use social media monitoring to measure their own impact in the digital world; to identify trends, communicate with customers, and to take control of the narrative surrounding a person or organisation. This not only helps in communicating with existing customers and identifying new ones, it also facilitates the identification of influencers who can help spread your message (which is not unlike what the purveyors of disinformation do). Overall, social media monitoring can be used to track the reach, influence, engagement, resonance, and reaction to a message/activity, and to analyse followers, mentions, and more to accrue information that will ease making positive communications decisions for an organisation and its mission.

CHALLENGES TO MONITORING FOR DISINFORMATION

There are many challenges that make monitoring social media for disinformation difficult. This section discusses a few of the more prominent challenges related to:

DATA VOLUME	Social media monitoring can yield so much data it becomes hard to proces
MANY CHANNELS	Many different platforms make encompassing monitoring difficult
MULTIPLE TOOLS	Different tools do different things and finding a tool or an API that will let you access the information you are interested in can be a challenge.
DIFFERING RULES	Different platforms provide access to different data.
VARYING LANGUAGES	Disinformation can come in any language and monitoring across multiple languages is both difficult and costly.
VARYING DEFINITIONS	When monitoring for disinformation, it may not be immediately obvious what you are in fact looking for.
NOT EVERYTHING IS TEXT BASED	Monitoring for images, speech, and other types of content is significantly more difficult than monitoring for text or specific users.

Challenges when first beginning to monitor are most often due to the volume of data. For example, being able to harvest thousands of tweets in real time might be deemed a success but sorting through your data in a manner that best fits your mission may present some obstacles. In fact, this problem has been a key factor in developing algorithms for topic discovery and event detection.[72]

As we have discussed, monitoring social media for disinformation is no easy task. Knowing where to begin with a tool requires significant work beforehand in order to get the most out of your monitoring. There is seemingly endless number of channels where information can be spread to users. You must decide which of these channels is right for your mission, depending on the population you wish to monitor or the types of information you would like to keep track of. This is important, as the conversations on Twitter differ from those on Instagram, Facebook, WhatsApp, Reddit, and 4Chan, as do the ages of users, their geographic locations, and the overall missions of the different platforms (see Chapter 4 for a detailed breakdown in how to design your monitoring strategy).

Before beginning to monitor, your biggest challenge may be deciding on the best tool for the job depending on the existing capabilities of your organisation. For some, a third-party tool is vital. For those with more advanced technological capabilities, using an API or a web scraper works best. But remember, the platform you want to monitor may not allow access to their API, so you would have to find a different method.

Identifying who, what, and how to monitor will depend on what you want to find out. The procedure for measuring the impact of a message will be different from that for discovering the source and spread of a message. Attribution is one of the top challenges in identifying disinformation via monitoring. It is not enough to identify what the message is, why it is being spread, and which networks are spreading it; you also want to identify the original source. However, in many cases attribution can be nearly impossible. Attribution is particularly important in cases of government-sponsored disinformation, where it might lead to international action, sanctions, or issues between countries or a population (more on attribution below).

Another challenge can be the platform itself. Twitter, for example, is far more open than Facebook. Facebook does not grant access to all public content in one go, and, if monitoring, one will likely need to customise searches and limit oneself to pages and groups. As a user, you constantly have to navigate restriction and limitation imposed by the platforms.

Language can also be a challenge when attempting to monitor social media. After all, if your goal is to identify, for example, anti-NATO messages in social media, they may well be in languages you may not speak or have the ability to translate. After all your hard work in identifying certain hashtags or influence networks, you might be left with unusable data. Having a plan for dealing with foreign languages is key, if you are monitoring internationally or in a multilingual environment.

Disinformation is created with the intent to deceive, but those who spread it may genuinely believe the content to be true in which case they would be spreaders of misinformation. You must decide what you or your organisation define as disinformation before proceeding to track it down. This will not only help speed up your monitoring and queries, yielding better results, but having a clear focus will lead to a better analysis.

The challenge of finding and monitoring images and other types of content that are not text based is also currently an issue, as there is no ready-made solution for tracking social media images, memes, or visual content on any platform. Much of the work on image tracking has been done manually. While there are some methods that allow for identifying brands or logos in shared images, detecting edited or fake images with one search, for example, is not currently feasible with conventional methods.

LIMITS OF SOCIAL MEDIA MONITORING

In addition to the challenges identified above, there are also limiting factors that constrain social media monitoring. Some of these, such as social media blind spots, are inherent to the nature of social media and its technical infrastructure. Others, like issues of attribution and legal constraints, relate to societal norms and regulatory considerations. Regardless, being aware of the factors that limit you monitoring is essential.

Asking the wrong questions

If you do not know what to ask, you will not find what you are looking for. Much as a poor search on Google will not yield any useful information, asking the wrong questions in your social media monitoring will not bring about the desired results and may end up adding to your workload. Before beginning, it is necessary that you have a clear focus in mind; formulate your questions unambiguously so you get answers you need (see Chapter "Methods for Monitoring").

Social media blind spots

Certain aspects of social media monitoring can confront you with frustrating blind spots. When looking for new or trending information, it can be challenging to know where to begin. Keyword- or hashtag-based monitoring assumes manual entry of terms, which you may not already know, and restricts your ability to identify what you are looking for—namely, pieces of disinformation that use precisely the new keywords and hashtags you are searching for. By solely monitoring known keywords or hashtags, your results are likely to be incomplete. Similarly, information may be outside of your research, locked away in closed groups or private channels, leading to a situation where you only see fragments of the bigger picture.

Content and users also travel across platforms. Monitoring one or two platforms may be sufficient to get an idea of what is going on, but disinformation is a complex phenomenon and what happens on one platform is not necessarily reflective of what is going on elsewhere. For this reason, it is useful to have a broad awareness of the social media environment (see chapter "Social Media").

Finally, you may want to monitor sentiment, or the reception that a piece of information receives, which is also notoriously difficult. This is partly because social media only indicate the sentiments of people who chose to engage – the sentiments of the silent majority are not presented – which creates a massive blind spot for disinformation research. Social dynamics reinforce this issue. For example, people with negative sentiments chose to express their opinions on social media to a higher degree than those with positive sentiments, and there are therefore more negative posts, in general, than positive.[73] This can lead to both bandwagon effects (where certain opinions are reinforced) and a spiral of silence (where contrasting opinions are ostracized).[74] This makes assessment of impact and reach of disinformation very difficult. It could be that most people who saw a post had a neutral to positive reaction and did not feel moved to respond while a small minority of vocal users chose to engage. Lack of complete information regarding user sentiment constitutes another blind spot.

Attribution

Attribution, or discovering the source of disinformation or content on social media in general, can be a difficult task, especially if faced with astroturfing where disinformation operators pretend to represent a grass-root movement.[75] Knowing the originator or instigator of disinformation is vital for measuring the scope or success of a certain message, as well as for determining its impact on society.

This issue came into play in the 2017 United States congressional hearing regarding infiltration of the 2106 election by Russia's Internet Research Agency; much of the evidence was rejected due to attribution challenges.[76] Since then, social media platforms have increased the transparency of their targeted advertisements, identifying those funding the advertisements, their affiliation, and more. However, the problem of attribution will only continue to grow as creators and disseminators of disinformation exploit loopholes and find ways to avoid identifying themselves. Social media platforms have also recently begun marking government-affiliated news sources, so readers of posts are aware of the originators' affiliations. You may discover a whole trove of what you think is state-sponsored disinformation but be unable to attribute it with confidence.

Ethical and legal constraints

While social media continues to evolve rapidly, watchers concerned with ethics and legal aspects struggle to keep up.[77] When dealing with personal data or information that may be of a sensitive nature, such considerations are crucial. Prior to embarking on a social media monitoring project, it is important first to review the ethical guidelines of your target platform. Most platforms have terms and conditions for platform users and for third parties who extract data. Is the data you wish to access truly publicly available? Is the data in a closed group that requires approval for access? Are the legal regulations in your country of operation conducive to this sort of monitoring and research?

What is within your scope? Some nations, such as the United States, protect hate speech and most forms of speech, including many forms of disinformation. However, some European countries now have laws allowing them to ban hate speech and to enact punishment for wilful spread of false information. Thailand, Indonesia, and Italy have enacted enforceable laws allowing the police to pursue legal action.[78]

A few guiding principles for ethical social media monitoring to keep in mind as you organise your process include:

- Assess the vulnerability of the audience or community you wish to monitor.

- As social media data generally involves real people (ignoring bots for the moment), adhere to the principles of human subject research, especially if the data can lead to the identification of private individuals.

- Carefully weigh the benefits of the research with potential privacy issues and keep in mind the rights of the subject. An example of this is the identification of populations most likely to engage in certain narratives or disinformation, which might lead to discrimination against that demographic.

- Ensure that any information you collect or save is done so in accordance with relevant legal frameworks, such as data protection and personal integrity.

METHODS FOR MONITORING

3. METHODS FOR MONITORING

Monitoring social media is not easy. It is impossible to collect every piece of relevant information and make sense of it. To break down, sort through, and process the abundance of information available you need a strategy for monitoring—a method for each challenge.

The methods you choose will depend on your goals for monitoring social media. If you are a journalist investigating disinformation campaigns, you will need to focus on verification of information and sources. If you work with open-source intelligence, you may be focusing on attribution instead. If you are monitoring brand engagement for your organisation, your will likely be interested in identifying evidence of engagement and consumer satisfaction.

Regardless of the tasks you have set, for the methods you choose needs to be effective, they must:[79]

1. Identify the problem you are interested in
2. Define and operationalise terms and concepts
3. Develop a plan for collecting information
4. Establish a mode of analysis
5. Create a structure for conducting the work and making use of the results

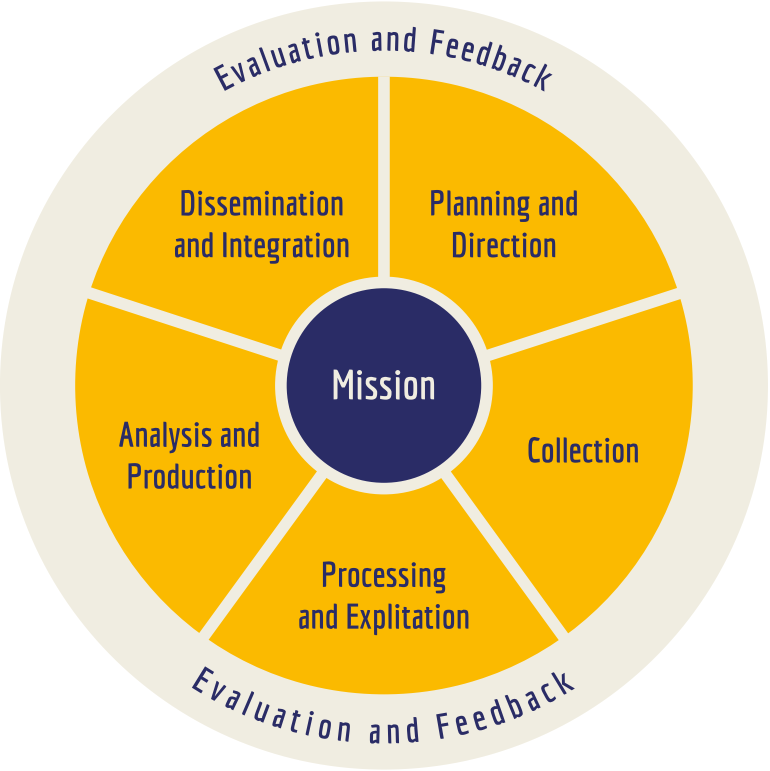
Developing a robust method for monitoring social media will not only make the monitoring easier, it will also be crucial for selecting appropriate tools (see chapter 5). First determine your goals, then select the best tools for the job. This may sound like common sense, but it bears repeating as there are many 'bright, shiny' social media monitoring tools on the market and it is easy to get carried away.

Regardless of how you design your monitoring method, it will be an iterative process. Unless you have a very narrow and specific task, your monitoring will involve the constant collection and processing of new information, and your operation will continuously evolve. For this reason, choosing a flexible and adaptable method is beneficial.[80]

If you have experience with academic research, the process of designing a research method will be familiar to you. If not, the more or less universal process of intelligence gathering, known as the intelligence cycle, will provide practical guidelines for designing your method and executing your monitoring.

THE INTELLIGENCE CYCLE

The intelligence cycle is a basic model of the analytical process used in the intelligence community to collect, process, and use information. It is by no means the perfect model for every task,[81] but it serves as a useful starting point for social media monitoring.



The intelligence process. Source: Joint Intelligence / Joint Publication 2-0 (Joint Chiefs of Staff)

The intelligence cycle is a well-established concept that has been written about at length elsewhere.[82] For the purposes of this paper, we will briefly highlight the most important considerations for each step of the process.

The intelligence cycle consists of five stages:[83]

1. Direction
2. Collection
3. Processing
4. Analysis
5. Dissemination

The intelligence cycle is a cycle for a reason—the process of establishing a direction, collecting information, analysing it, and producing meaningful outputs is a continuous one.[84] The results of each monitoring cycle provide feedback that drives new information needs and leads to fine-tuning for the next cycle.

Direction

The first step in the cycle is direction. This is the most comprehensive step as it encompasses the planning and management decisions for the entire collection and analysis effort. The point of departure for this stage is a clear information requirement—that is, an explicitly stated understanding of what you are looking for. In intelligence organisations, this usually comes in the form of an intelligence requirement from a commander or another organisation, but you can also formulate your own information needs if you do not have a clear request in your organisation.

Information needs

Defining your information needs is crucial to establishing the scope of your monitoring. What is it that you are interested in? Are you looking for disinformation related to a specific topic? Are you interested in the activities of a specific interest group or movement, or perhaps a specific platform? Which audience are you interested in? Answer these basic questions fully and clearly before you begin.

What are you interested in?

What is the core issue your monitoring will address?

What type of information are you looking for?

What are your core research questions?

What is the final product you are looking to produce?

Definition of concepts and terms

It is equally important to define the concepts with which you will operate. If you want to monitor social media for disinformation, what is your working definition of disinformation? What falls outside of your definition? Having a clear understanding of the concepts and terms you are using is fundamental to the success of your efforts.

If you conduct your monitoring on behalf of an organisation or agency, your definitions need to be aligned with the definitions accepted by your employer. There may be legal aspects to consider at this stage to ensure that your monitoring efforts are in line with your mandate.

Operationalisation

Operationalisation is the process of turning your working definitions into measurable factors and indicators.[85] Operationalisation is crucial for any type of research. To find what you are looking for, you need to know which tell-tale signs to look out for.

Investing some extra time in operationalisation will pay large dividends down the line by increasing the quality and accuracy of your results. Well-executed operationalisation will also increase the transparency of your research process and ensure you don’t end up collecting information you do not need or should not have.

Operationalising disinformation and information influence activities

The report *Countering Information Influence Activities* (Pamment et al. 2018) offers an example of how a complex definition can be operationalised by setting up four diagnostic criteria derived from the Swedish understanding of ‘information influence’:

DECEPTION There must be an element of deception, i.e. something factually incorrect or misleading.

INTENTION There must be an intent to deceive or mislead.

DISRUPTION The activity must have a disruptive effect.

INTERFERENCE There must be an element of interference, i.e. foreign involvement, or proxies.

These four criteria can serve as a basis for developing indicators of what to look for when monitoring social media and the online information environment.

Collection Plan

Information needs, definitions, and indicators comprise your collection plan—your strategy for using available resources to access the information you require.[86]

A collection plan does not need to follow a specific format, but it should contain detailed, transparent information about your priorities and about how you will collect your data.

Consider what information is most useful to you. Are you interested in the content of a post, the account that posted it, the accounts that engaged with a post, the network in which an account is situated, or all of the above? Thy type of data you collect will depend on your needs and interests.

Collection plan

Make sure your collection plan answers the following questions:

- ? What are your information needs?
- ? What indicators are you monitoring for?
- ? When will you begin and end the monitoring?
- ? At what intervals will you monitor?
- ? Which audiences will you monitor?
- ? Which platforms will you monitor?
- ? Which tools, services, and sources will you use?
- ? How will you prioritise tasks?
- ? How will you log/store/save your data?

The 3 M's

In their excellent guide to social media monitoring in the context of elections *Democracy Reporting International* (DRI) suggests three aspects to study—the 3 M's.

Message: The content of the message.

Messenger: The sender of the message.

Messaging: The distribution of the message.

This simple typology provides a clear structure for a data collection plan and is useful to provide direction to your analysis.[87]

Collection

The activities specified in your collection plan are executed in the collection phase as you monitor systematically to gather data in support of your information needs. It is likely that you will need to use a combination of collection tools and techniques to gather sufficient data.

The data you collect can come in many shapes and forms. It is rare for a single type of data to answer a complex question. If you are tracking disinformation online, you may need to collect data regarding content, sentiment, geolocation, networks, etc. to understand how a piece of disinformation is disseminated and what impact it is having.

Collection normally occurs continuously or at set intervals; in both cases you will move back and forth between collection and processing/analysis. You do not need to 'finish' your collection to progress to the next step.

Collecting information

When collecting data, consider the following:

- Document/record your findings
- Collect from multiple sources
- Stick to the collection plan
- Ensure consistency over time and across users

Platform specific collection methodologies from DRI

Democracy Reporting International (DRI) has constructed a Digital Democracy Monitoring toolkit which offers custom method suggestions adapted for specific platforms. Their guide is a valuable resource for designing your collection plan.

Processing

The data you collect will be 'raw' or 'unfiltered'. For example, if you are monitoring Twitter by scraping a set of hashtags for certain keywords over a set period of time, you will end up with a raw directory of tweets that may or may not be relevant for your analysis. This data needs to be processed, organised, and synthesised for analysis to be meaningful. [87]

At the processing stage, you convert raw data into a useful product by, for example, structuring information, ordering data, translating language, ensuring data quality, removing outliers, decrypting information, removing noise, or simply collating data into a single dataset.[88]

Processing can be done in parallel to the collection phase if you are monitoring continuously or simply wish to speed up the process.

Analysis

Analysis transforms interesting information into actionable information.[89] You method of analysis will depend on the type of data you have collected—if you have found an interesting pattern of accounts that spread disinformation, you may use network analysis to understand their relationship, but if you are interested in the discourse of a particular set of users you may want to use content analysis to produce the desired results.

There is no one single mode of analysis that fits every task. Analysis generally involves integrating and evaluating processed data to discern patterns, draw conclusions, and integrate information into a wider context.[90]

Types of analysis for disinformation and information influence activities

Many different types of analytical models are useful for social media monitoring. Some examples that have been used for identifying and analysing disinformation and information influence activities include:

NETWORK ANALYSIS	depicting relations among actors to analyse social structures[91]
SENTIMENT ANALYSIS	analysing subjective information by determining its sentiment[92]
CONTENT ANALYSIS	assessing patterns of communication in message content[93]
TEXT ANALYSIS	detecting and interpreting trends and patterns in text[94]
LINGUISTIC ANALYSIS	studying the form, meaning, and context of language used[95]
STATISTICAL ANALYSIS	using quantitative data to draw inferences[96]
GEOLOCATION ANALYSIS	assessing the geospatial data of social media posts[97]
TREND ANALYSIS	searching for patterns and trends in available data[98]

Dissemination

The term 'dissemination' has a particular meaning in the context of intelligence analysis—it is the stage at which an intelligence product is delivered to the commander or policy maker who requested it. For you it may mean something different, but always keep in mind that the output of your monitoring should be useful in the end.

Once you have collected, processed, and analysed your information and data, you should have an actionable product that can be used to achieve the goal of your monitoring operation. To ensure that the results of your efforts are useful, align your outputs with your organisational needs and with the formats and templates your organisation uses for other information.

Feedback

While not a specific step of the intelligence cycle, continuous feedback is crucial for aligning your monitoring work with your information needs as conditions change over time. Feedback should occur throughout every stage of the process to ensure flexibility and adaptability as conditions and circumstances change.

Other considerations

When designing your monitoring method, the following considerations are also worth keeping in mind.

Check your biases

Any type of research runs the risk of bias—predispositions, inclinations, assumptions, and prejudices affect results if we aren't careful. When designing your monitoring method, check you monitoring system for biases to ensure that you are not drawing the wrong conclusions from your data. While you may never be able to design a perfectly unbiased monitoring system, being aware of your biases is the first step of mitigating them.

When it comes to monitoring social media, watch out for the following three common biases as a matter of course. First, confirmation bias—or interpreting information as confirming your pre-existing beliefs or your hypothesis.[99] This is one of the most common research biases and in our field can take the form of assuming that every post conforming to a narrative associated with an influence campaign is a piece of disinformation. Minimising confirmation bias requires continual re-evaluation of evidence and assumptions to make sure they hold up to scrutiny.

A second bias to watch out for, and one that is particularly troublesome with regard to social media, is selection bias—when an unrepresentative sample is assumed to represent an entire population.[100] It seems intuitive that social media, with its billions of users, provides data that is representative

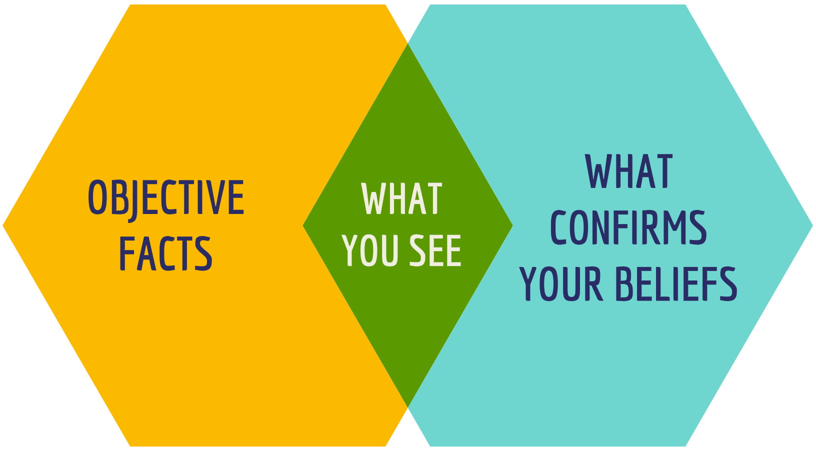


Figure 4 An Illustration of confirmation bias

of the population, but variations in media consumption patterns and other offline factors can easily give rise to selection bias. Analysts must take care to understand the audiences they are monitoring and whom they represent when interpreting their results. Sentiment analysis is particularly vulnerable to selection bias, as sentiments expressed on social media do not capture the opinions of silent users who read, but do not comment on or write public posts.[101]

The third bias to keep in mind is data bias—a bias resulting from the ways social media platforms make data available for research.[102] When collecting data from a social media platform, either manually or via a tool or database provided by the social media platform, you will rarely get all of the available data. Most likely you will get a snapshot of a wider data set, or a curated data set. And even if you are able to receive a complete data set, it may be biased due to structural characteristics of the system that produced or collected the data.[103] Of course, you can draw valid conclusions from imperfect data, but attention must always be paid to the quality of data you are working with.

Structured Analytic Techniques

A tried and tested way of working consciously to minimise bias is to apply Structured Analytic Techniques (SATs) to your work. These techniques offer standardised methods for mitigating cognitive pitfalls and are often used by intelligence analysts when assessing information. These techniques are also useful when assessing disinformation on social media.[104]

For more information on SATs, check out:

- CIA (2009) *A Tradecraft primer: Structured Analytic Techniques for Improving Intelligence Analysis*[105]
- Heuer & Phearson (2010) *Structured Analytic Techniques for Intelligence Analysts*[106]
- GCS (2019) RESIST: *Counter Disinformation Toolkit*[107]

Manage uncertainty

There will always be an element of uncertainty to your monitoring and resulting analytical assessments. If you spot a network of accounts operating in a coordinated fashion where many of the accounts are anonymous and were recently created, it is fair to assume that you have discovered a bot network—but with the information you have, this will be a qualified assessment at best. Without technical and data forensic information, you may not be able to confirm with certainty that it is, in fact, a bot network.

Multiple techniques for managing uncertainty are currently being practiced by researchers and decision-makers. Without going into detail, the point here is to always be aware of uncertainties and communicate the degree to which you are confident in your results and conclusions.

Ensure transparency and accountability

Finally, you should always consider the degree to which your monitoring is designed and conducted in a transparent and accountable fashion. From both a democratic and a legal perspective, it is vital to ensure that your actions are open to oversight and hold up under scrutiny.

TOOLS FOR MONITORING

> I know everyone wants a tool in which they can just plug in a website, social media account, or post and have automated program tell you with certainty if it's disinformation. But the reality is, there's always going to be a need for some human analysis on top of it.
>
> Cindy Otis, Twitter, 19 Jan 2020

4. TOOLS FOR MONITORING

As we have established, social media monitoring involves much more than booting up a piece of software that will tell you what's happening on social media. That being said, there are many commercial products and tools that can assist you in collecting, processing, and presenting the complex information you will glean from your monitoring.

This chapter provides an entry point into tools and services for social media monitoring. It does not give an exhaustive review of all the available tools, but it does provide an overview of some of the current tools which can be used for monitoring purposes.

An up-to-date list of every tool available is difficult to maintain as both the social media platforms and the technology for monitoring them develop continuously. The selection presented here was chosen based on consultations with subject-matter experts who work with social media monitoring related to disinformation and information influence activities and were kind enough to share their preferences and recommendations. We then tested the tools ourselves to get a sense of their comparative benefits and drawbacks. The selection is therefore by no means exhaustive or definitive but gives a snapshot of some of the tools available on the market.

Some additional resources on monitoring tools are provided in the box on the next page.

Additional resources regarding reviews and recommendations for social media monitoring tools		
Social Media Monitoring Tools and Services Report	Ideya	Reviews for hundreds of social media monitoring services
Guide for Civil Society on Monitoring Social Media During Elections	Democracy Reporting International	Guide to social media monitoring during elections designed for civil society organisations
Fighting Disinformation Online: A database of web tools	RAND Corp.	Database of web tools for fighting disinformation
Disinfo Cloud	Disinfo Cloud / Global Engagement Center	Inventory of hundreds of vetted companies and tools for social media monitoring
First Draft—Basic Toolkit	First Draft	Toolkit of free monitoring and news-gathering tools
Data Analytics for Social Media Monitoring	National Democratic Institute	Guide to social media analytics for activists and researchers

SOCIAL MEDIA MONITORING TOOLS

While there are simple monitoring tools and supporting software available online for free (see next section), most monitoring suites or advanced tools are subscription-based services, with prices ranging from a couple hundred USD per year, to thousands of USD per year. Generally, more expensive tools offer more comprehensive and useful services, but this is not always the case. Different tools offer different solutions, and the optimal tool will depend on your needs.

Below we provide reviews of ten social media monitoring tools identified as useful for monitoring for disinformation and information influence activities by our consulted experts and practitioners.

Name	Focus	Media	Languages	Price range ($-$$$)[108]	Access to API	Historical data
Brandwatch	Social listening, trends identification, image analysis, sentiment scoring, influencer identification	Full data firehoses from Twitter, Reddit, and Tumblr; other sources include social networks, blogs, forums, news sites, review sites, and video sites	Collects data written in any language; sentiment analysis in 44 languages	$$$	Yes	Yes (The Consumer Research data archive contains 1.2T+ documents/posts, with historical data going back to 2008)
BuzzSumo	Content discovery, influencers identification, competitor analysis, alerts	The web in general, Facebook, Twitter, YouTube	Search supported in all languages; sentiment analysis is not available	$–$$	Yes (Account API and Search API)	Yes (depending on the plan; with Pro and Plus—1 year, Large—2 years, Enterprise—5 years)
Hootsuite	Managing social media accounts	Various networks, including Twitter, Instagram, Facebook, LinkedIn, Reddit, and others; blogs and forums as well	Around 50 languages for mentions; sentiment in over 20 languages	$–$$	Yes (Publishing API, User Management API and others)	With Brandwatch integration—yes; otherwise limited historical data available on your own accounts
Lexalytics	Text documents analysis	n/a (does not source data)	Over 25 languages (available features differ by language)	$$$	Yes (Semantria API)	n/a (does not source data)
Meltwater	Social listening, social analytics, Twitter insights, media intelligence	Twitter, Facebook, news, blogs, forums, Reddit, YouTube and others	Sources in over 80 languages; sentiment analysis in 16 of those	$$$	Yes (Export API and Analytics API)	Yes (depending on the source itself)
Nexalogy	Social listening, actionable intelligence	Twitter, blogs, Facebook and custom RSS feeds	n/a	$$	Yes	n/a
Pulsar	Social listening, audience intelligence, trends tracking	Facebook public pages, Twitter, Instagram, YouTube, Tumblr, news, blogs, forums and others	Sources in over 180 languages; sentiment analysis in 26 of those	$$	Yes (the Pulsar API can only be used with a token)	Yes (depending on the source itself)

Name	Focus	Media	Languages	Price range ($-$$$)[108]	Access to API	Historical data
Sprinklr	Social listening, management of social media accounts	25 social platforms (including Instagram, Facebook, Twitter and LinkedIn), web sources (blogs, new sites etc.) and 11 messaging channels	Supports translations in over 80 languages and sentiment scoring in 17 languages	$$$	Yes	Yes (depending on the source itself)
Sprout Social	Managing social media accounts	Twitter for Professional plan; multiple sources (Facebook, Instagram, YouTube, Reddit, Tumblr, the web) for Advanced plan	Search in over 20 languages	$–$$	No	Yes (depending on the source itself)
Talkwalker	Social listening, audience intelligence, trends tracking	Various social networks (Facebook, Instagram, Reddit), including full firehose access to Twitter, news sites, blogs, forums, print, broadcast (but depends on a plan)	Monitoring in 187 languages; sentiments analysis in 25 of those	$$$	Yes (a significant number of different APIs available)	Yes (2 years of it available as an add-on for all plans)

Brandwatch

Brandwatch is a consumer intelligence service which offers social media monitoring including blogs, news sites, forums, video services and social networks. In 2018, a merger between Brandwatch and Crimson Hexagon was announced, with a promise to combine the best features of both products into one platform—Crimson's historical data, machine learning, and advanced AI, and Brandwatch's superior user interface and its fast and flexible data handling.[109] With an extensive set of features and an access to a large dataset, Brandwatch's platform is consistently recommended by various specialists as the best tool for social listening, including for purposes of detecting and tracking disinformation online. The tool offers both text and image search functions and includes visualisation and sentiment analysis features. This greater functionality and richness of data, however, have their downside—it is notably more expensive than most of its competitors and can be difficult to navigate, even for people who have previous experience with a different listening platform.

Pros	Cons
+ High customisation and flexibility	- High cost
+ Great number of data sources	- Steep learning curve
+ Multiplicity of related products and add-on services	
+ Allows for very specific queries that are easy to create	
+ **Generally considered best choice in the field**	

BuzzSumo

BuzzSumo, a software solution provided by Brandwatch, offers a range of monitoring tools; their sophistication and flexibility vary by payment plan. BuzzSumo allows users to track popular content by topic, user, or website. It scans social sites to see what people are discussing. While it does not allow for social listening in the strictest sense, BuzzSumo provides valuable insights into the top and trending posts on various social media platforms and websites in general, with up to five years of historical data available. To discover this top content, both keywords and pages can be searched. BuzzSumo data can also be accessed through two REST APIs. BuzzSumo analyses posts and can provide, for example, information about the length and type of the most popular posts, the optimal time for posting, etc. In addition, BuzzSumo offers a comprehensive way of setting up alerts for mentions of specific keywords. BuzzSumo can be a useful tool for detecting disinformation online, especially when combined with the complementary capabilities of other platforms—for example, Sprout Social or Lexalytics.

Pros	Cons
+ Easy to navigate + Great for discovering top content quickly (e.g. top influencers, top articles, trending topics) + Relatively inexpensive (especially the two most basic plans); transparent pricing + Offers a free 7-day trial period for three of their plans	- Rudimentary Boolean search - Simple visuals - Sometimes irrelevant/duplicate results are returned - Access to some basic features (e.g. top author search) costs more

Hootsuite

As a social media management platform, Hootsuite does not exclusively focus on social listening, but offers it as a key feature. Its comprehensive description of features and how to use them (including the free online course 'Hootsuite Platform Training') makes this platform easy to use. There are two ways to conduct social listening using Hootsuite—basic listening can be done through Hootsuite itself, while a more sophisticated and in-depth listening platform can be accessed through Hootsuite Insights—a relatively new add-on powered by Brandwatch (available only to customers paying for the two most expensive plans). Hootsuite Insights supports around 50 languages, and provides sentiment analysis in over 20 of them, including Arabic, Chinese, and Russian. Since Hootsuite's APIs are open source, applications developed separately can interact with the platform and access its data. However, at present, Hootsuite does not offer a Hootsuite Insights-specific API.

Pros	Cons
+ Offers a limited free version	- Best suited for simple social media management
+ Brandwatch integration	- Lack of analytics (apart from Hootsuite Insights)
+ Easily accessed training materials; freely available explanatory videos	
+ Relatively inexpensive; transparent pricing	
+ Offers a 30-day free trial period (billing information required, however)	
+ Great for managing multiple accounts	

Lexalytics

The Lexalytics Intelligence Platform focuses on text analysis, which sets it apart from other platforms on the list. Even though Lexalytics does not source any social media data, its various solutions can be used to fill in the gaps left when applying other tools. The platform provides a semi-customizable solution, making it possible to design an individualized for your monitoring. Lexalytics' natural language processing (NLP) APIs and Semantria Storage and Visualization tool are such useful complements to the process of social listening that some outside companies have integrated these NLP APIs into their own social listening products. The Semantria Storage and Visualization tool stores and analyses text documents, breaking them down by topic, theme, or category to provide summaries, and prevalent sentiment; you can build dashboards with a multitude of visualisations. Lexalytics' text analysis features are offered in over 20 languages; however, not all of these features are supported in every language—currently only texts in English can be analysed fully.

Pros	Cons
+ A very specific focus on text analysis + Identifies themes, entities, and the sentiment associates with a text + Easy to integrate with other products + A demo version of the text analysis tool is available on the website	- Does not source social media data - Not much information available from Lexalytics or third parties on functionality, ease of use, or best features

Meltwater

Meltwater positions itself as a media intelligence company that provides monitoring for both traditional and social media. With its neat and easy-to-use interface and numerous social listening capabilities, Meltwater's platform allows users to monitor and better understand online interactions in over 80 languages. Sentiment analysis available in 16 of those, but not for Russian-language sources. It is considered particularly useful for identifying influencers and opinion leaders. Meltwater's Export API and Analytics API make access to data collected for export purposes and integration with other systems fairly straightforward. The downside is its price; customers sometimes describe the platform as 'too expensive'.

Pros	Cons
+ Relatively easy to navigate; easily built dashboards	- High cost
+ Comprehensive influencer/opinion leader search	- Steep learning curve
+ Mobile app allows to access many of Meltwater's features, including saved searches, analytics, and interactive dashboards	- Sometimes returns non-relevant results
+ Both social and traditional media monitoring	

Nexalogy

A leading provider of social media research technology in Canada, Nexalogy offers a free version of its services as well as standard paid solutions. The free version allows users to receive a comprehensive data analysis of the user's own Twitter account, including, for instance, a network graph, timeline, top concepts, links, and hashtags. NexaIntelligence, its paid software solution, provides deep analysis of social media data (e.g. discovery of top content and terms, top actors and their interactions) with various useful, albeit not necessarily the most engaging, visualisations such as actor interactions maps (presenting conversation communities) and lexical maps (presenting the top 200 concepts and the relations between them). While Nexalogy is among the lesser-known providers of social media listening tools, it focuses on government research and non-profits and not advertising and marketing as most of its competitors do.

Pros	Cons
+ Offers a limited free version	- Not much information available from Nexalogy or third parties on its functionality, ease of use, or best features
+ Comprehensive visualisations (e.g. to better understand interactions between actors)	
+ Focus on services for public sector and governments	

Pulsar

Described as a social listening and audience intelligence platform, Pulsar offers four different solutions—Pulsar TRENDS, Pulsar TRAC, Pulsar CORE, and Pulsar RESEARCH. Pulsar TRENDS and Pulsar TRAC are best suited for detecting and tracking disinformation. While the TRENDS solution can trace the trajectory of engagement for trending topics over time and can show how conversations spread, TRAC is a social listening tool that can search in over 180 languages and analyse sentiments in 26 of those. Depending on the source, the Pulsar platform also collects historical data ('historics'); for Twitter, it goes as far back as 2006, while for the majority of other sources historical data is collected for the past 25 months. Pulsar's abundance of visualisations and other features can be overwhelming, especially for users with limited experience

Pros	Cons
+ Neat, user-friendly interface	- High cost but transparent pricing
+ Intuitive and diverse visualisations	- Can be overwhelming at first due to its many sophisticated features
+ Clearly categorised solutions for a variety of objectives	

Sprinklr

Sprinklr offers a unified platform—described by its VP of Marketing as 'the only platform that brings together capabilities for marketing, advertising, research, care and engagement in a single environment with a consistent user experience'[110] —that is primarily focused on two functions, listening and managing. A key advantage of Sprinklr as a management tools is that all team members have real-time access to the changes and decisions being made. As a listening tool, Sprinklr has access to a wide range of sources and data, which allows it to listen to online conversations (and track disinformation), even on platforms that are often overlooked, such as Sina Weibo. Moreover, Sprinklr not only allows users to listen to various text publications in many languages, it can also conduct text-based search for images on Twitter in a number of languages; the platform claims to be able to detect people, places, objects, and logos in images, which provides information on how they are used. However, like Brandwatch, Sprinklr is often described as expensive and having too steep a learning curve.

Pros	Cons
+ Integrations offered with a variety of other—mostly business-related—applications (e.g. ServiceNow)	- High cost
+ Mobile app that accesses many features, such as creating/reviewing listening and reporting dashboards, receiving mobile notifications, etc.	- Better suited to social media management
+ Intuitive and diverse visualisations	- Business-focused
+ Able to conduct text-based image searches on Twitter in a number of languages, and detect people, places, objects, and logos in images	

Sprout Social

Sprout Social is similar to Hootsuite in that both platforms are designed for managing various social media accounts. Social media listening is not a core feature—it is one of many features created to help users organise and keeping track of their own social media accounts. Listening is available only to customers paying for the Professional and Advanced plans; while more expensive than the company's Standard plan, they are slightly cheaper than the other solutions listed here. Despite the limited functionality of Sprout Social's listening feature compared to more powerful and social media listening-focused platforms, it allows for various types of keyword searching, including by hashtag, word and phrase, and user-mentions in over 20 languages. It has some useful filters for tweaking search queries, which make it easy to use for people with limited experience.

Pros	Cons
+ Easy to navigate + Relatively transparent pricing + Offers a free 30-day trial period for three of their plans + Useful and detailed query builder + Great for managing multiple accounts	- Best suited for simple social media management - Social listening available only with Professional and Advanced plans

Talkwalker

Like Nexalogy and Hootsuite, Talkwalker offers some simple search and analysis options for free—namely, Talkwalker alerts (similar to Google Alerts; the two alert systems can be used to complement each other, ensuring no trend or mention is missed) and Talkwalker's free social search function that monitors campaigns and hashtags and provides basic analysis. The paid version can track conversations and sentiment online in general, on social media, on TV and radio, and in print.

Talkwalker is particularly good at predicting trends and tracking and measuring them over time. A Hootsuite-Talkwalker integration feature is available via the Hootsuite App Directory. As is often the case, many of the most interesting features such as image recognition, historical data access, the AI engine, and advanced integration are available as add-ons at an additional cost, despite the paid solutions already being quite expensive.

Pros	Cons
+ Offers a limited free version with comprehensive visualisations + Supports integration with a variety of other tools (Hootsuite-Talkwalker integration available) + Customisable and flexible + Allows for very specific queries	- High cost - Steep learning curve

Supporting tools and services

As none of the tools listed in the section above—no matter how comprehensive or expensive—can provide a one-size-fits-all solution, it is important to always keep an eye on other tools that can fill the gaps in your monitoring, which of course depends on the goal of your project.

For example, when conducting social media monitoring using Meltwater or Pulsar, the free Google Chrome extension Bot Sentinel can be used to check the likelihood that accounts coming up as query results are in fact bots. Results from free/low-cost tools, apps, and extensions should always be treated with caution, but can provide a good starting point for the human analysis that is necessary in all cases. While many free tools seem to be offering a variety of useful features, in reality many of them are outdated or simply don't work as intended.

Below we list a number of supporting tools and services that can be useful for identifying and tracking disinformation on social media:

Name	Type	Description	Link
Bot Sentinel	Browser extension	Rates Twitter accounts on how likely they are to be bots and adds them to a database available to the general public. It uses machine learning and artificial intelligence.	https://chrome.google.com/webstore/detail/bot-sentinel/eadmnplpcakhnmjbaioehol-pakbknhgc
CrowdTangle	Browser extension	Checks how many times a link has been shared, by whom (if available, e.g. public profile), the total number of interactions, etc. on Facebook, Twitter, Reddit, and Instagram.	https://apps.crowdtangle.com/chrome-extension
FeedReflect	Browser extension	Makes posts on Twitter either more or less visible, reflecting their purported reliability.	https://chrome.google.com/webstore/detail/feedreflect/bigmeipgaifggglelcnpnp-baefimpooc?hl=en-US
InVID	Browser extension	Provides video and channel/user metadata (via analysis and metadata tools), identifies keyframes, and perform reverse image searches (via keyframes tool), zooms in on photos to study details (via magnifier tool), and more for videos uploaded from YouTube or public Facebook/Twitter pages.	https://chrome.google.com/webstore/detail/fake-news-debunker-by-inv/mhccpoaf-gdgbhnjfhkcmgkndkeenf-he?hl=en
Microsoft Newsguard	Browser extension	Qualifies and categorizes online news articles according to reliability using a journalist-based crowdsourcing mechanism.	https://microsoftedge.microsoft.com/addons/detail/newsguard/cgooaaonimepb-cidkhgmanahfbinpjdm

Name	Type	Description	Link
GDELT Project	Website	The GDELT (Global Database of Events, Language and Tone) Project searches press and media reporting by keywords, country, source, domain, and other filters and provides basic analysis.	https://gdelt.github.io
GLTR	Website	Predicts how likely it is that a certain text was generated automatically by employing the same language models that are used to generate fake texts.	http://gltr.io/
Google Alerts	Website	Notifies the user when new content related to their search becomes available.	https://www.google.com/alerts
Google Trends	Website	Allows the user to see trending searches and conduct their own trend searches by theme, country, or source-type. Provides breakdown by region and sub-region and lists related searches.	https://trends.google.com/trends
Deepware Scanner	Website	Uses Artificial intelligence to detect fake YouTube videos.	https://deepware.ai/deepware-scanner/
Snap Map	Website	Allows users to discover and browse Snapchat content based on location. Users can either explore or find locations on a map and see recent posts made in that location.	https://map.snapchat.com
TinEye	Website	Tracks whether and where the images a user uploads, pastes, or searches via the URL on TinEye's website have appeared online.	https://tineye.com
Tweetdeck	Website	Manages multiple Twitter accounts, tracks trending topics in any country, performs searches by keyword or phrase and filters the search by location, language, engagement, etc.	https://tweetdeck.twitter.com/
SearchUsers.com	Website	Allows users to search for Instagram accounts by name or username without requiring an Instagram account to do so.	https://searchusers.com
Wopita	Website	Allows users to search for Instagram users and hashtags without requiring an Instagram account. The number of photos that can be viewed is limited.	https://wopita.com/
TrustServista	Service and browser extension	An AI-powered content verification algorithm that provides three types of solutions—web dashboard, TrustServista REST API, and Google Chrome Extension.	https://www.trustservista.com

Bespoke solutions

There are also alternative ways of monitoring social media. It is not uncommon, for example, for organizations to develop in-house bot detection or video analysis software. Access to APIs provided by the majority of the platforms makes it possible to integrate those capabilities into the products developed in-house. This approach ensures that the end solution is best suited for the specific aims of the organization.

However, both detecting and tracking disinformation online is a time-consuming endeavour, especially if developing a unique bespoke product. Outsourcing these tasks to external data scientists, analysts, or consultancy agencies can prove a practical alternative.

It is always worth considering whether investing in an existing tool is the best use of available resources. Sometimes alternative solutions—developing an in-house tool or outsourcing the process of social listening altogether—might prove a better use of time and funds.

Examples of bespoke solutions
Many companies and actors that provide solutions for bespoke social media monitoring. Below are three examples of such services that we have sampled:

BBC Monitoring
As a specialist unit within BBC News, BBC Monitoring has a dedicated disinformation team tasked with identifying, collating, and investigating examples of misleading reporting and manipulation. The NATO StratCom COE collaborated with BBC Monitoring to study the techniques and tactics of social media manipulation in the report Malicious Use of Social Media: Case studies from BBC Monitoring (2018).

Storyzy
Storyzy is an online media intelligence company that offers customized monitoring at the intersection of traditional media and social media. With a database of thousands of categorized sources and websites, the service makes it easy to visualize disinformation networks and information flows online.

Singularex
Singularex is a social media intelligence and analytics company that provides bespoke solutions in the realm of data science and social media. The NATO StratCom COE has used Singularex services on many occasions, most notably for the report Falling Behind: How Social Media Companies are Failing to Combat Inauthentic Behaviour Online (2019).

What tools will not do for you

There are many things that tools can do. However, there are also many things that they cannot do for you. Importantly, they cannot with certainty detect disinformation on their own; the need for human analysis—either in-house or outsourced (some platforms even offer analysis services alongside other features)—is a crucial point that should not be overlooked.

The results of social media monitoring depend entirely on the queries built by the user. Some platforms have more precise and comprehensive query-builders that allow for narrower and more specific searches, while others have quick and easy ones that give less precise results. Some platforms also have access to more data sources than others, returning greater volumes of relevant results. Despite these significant differences, results returned still depend primarily on how a query is defined by the user. A good understanding of Boolean searching makes properly defining queries easier.

Another thing tools cannot do is determine with certainty whether or not detected online activity is authentic or inauthentic. Even the various bot detectors available—both paid and free—should be used with caution and in concert with human analysis. The various visualisations produced by the tools, however, can aid you in determining whether activity is authentic, for example, by mapping how information spreads and how the accounts spreading it interact.

Anyone engaged in social media monitoring would like to have neat and definite answer to what is and is not disinformation, and who and how why it is spread. To get the best possible results, it is crucial to understand the limitations of social media monitoring tools. None of them can do everything on their own. Some tools produce better results than others, but they all require humans to define tasks, to supervise, to complement with additional tools, and to interpret the results.

Endnotes

1. *Allied Joint Doctrine for Psychological Operations*, (NATO/UK Ministry of Defence joint publication, September 2014). [Retrieved 8 October 2020].
2. C. Shao, G. L. Ciampaglia, A. Flammini & F. Menczer, 'Hoaxy: A Platform for Tracking Online Misinformation', WWW '16 Companion: Proceedings of the 25th International Conference Companion on World Wide Web, April 2016, p. 745–50.
3. T. Nissen, *The Weaponization of Social Media,* (Royal Danish Defense College, March 2015). [Retrieved 8 October 2020].
4. NATO defines 'hybrid threat' as a 'type of threat that combines conventional, irregular and asymmetric activities in time and space'. See: AAP-06 NATO Glossary of Terms and Definitions, (NATO Standardization Office (NSO), 2019). [Retrieved 8 October 2020].
5. G. F. Treverton & R. Miles, *Social Media and Intelligence,* (Swedish National Defence College, 2014).
6. A. Shahbaz & A. Funk, 'The Crisis of Social Media', *Freedom House*, 2019. [Retrieved 8 October 2020].
7. J. Pamment, H. Nothhaft, H. Agardh-Twetman& A. Fjällhed, *Countering Information Influence Activities: The State of the Art* (Department of Strategic Communication, Lund University, 1 July 2018). [Retrieved 8 October 2020].
8. *Social Media Monitoring Tools and Services Report Public Excerpts 2018: Analysis and Elaborate Profiles of More than 150 Social Technologies & Services Worldwide*, Ideya Market Report, 9th Edition, November 2018). [Retrieved 8 October 2020].
9. W. Marcellino, M. Smith, C. Paul & L. Skrabala, *Monitoring Social Media: Lessons for Future Department of Defense Social Media Analysis in Support of Information Operations*, RAND Corporation, 14 June 2017. [Retrieved 8 October 2020].
10. *Social Media Monitoring Tools and Services Report*, Ideya Market Report.
11. E. J. Briscoe, D. S. Appling & H. Hayes, 'Cues to Deception in Social Media Communications', 2014 47th Hawaii International Conference on System Sciences, p. 1435–43.
12. Pamment et al., *Countering Information Influence Activities*.
13. This was highlighted already in 2002 by Rathmell who argued for a new paradigm of information sharing between and within states to achieve a full intelligence picture in the online environment. See: Rathmell, 'Towards', *Intelligence and National Security* Volume 17 Issue 3(2002): 87–104;
14. A. M. Schejter & N. Tirosh, '"Seek the meek, seek the just": Social media and social justice', *Telecommunications Policy* Volume 39 Issue 9 (2015): 796–803.
15. J. H. Kietzmann, K. Hermkens, I. P. McCarthy & B. S. Silvestre, 'Social media? Get serious! Understanding the functional building blocks of social media' *Business Horizons,* Volume 54, Issue 3 (2011): 241–51; J. A. Obar & S. Wildman, 'Social media definition and the governance challenge: An introduction to the special issue', *Telecommunications Policy* Volume 39, Issue 9 (2015): 745–50.
16. N. Erragcha & H. Babay, 'Social Media, Marketing Practices, and Consumer Behavior' in N. M. Suki & N. M. Suki (eds), *Leveraging Consumer Behavior and Psychology in the Digital Economy Advances in Marketing, Customer Relationship Management, and E-Services,* (IGI Global, 2020) p. 27–45.
17. Kietzmann et al., 'Social media? Get serious!'.
18. Schejter & Tirosh, '"Seek the meek, seek the just"'
19. Shao et al., 'Hoaxy: A Platform for Tracking Online Misinformation'.
20. Alexa Website Traffic (2020). [Retrieved 8 October 2020]. Please note that Alexa's global internet engagement rankings measure web traffic, so for services that that are almost entirely app-based, such as Snapchat, this ranking does not fully capture its popularity.
21. R. Manokara & M. Paramonova, *Manipulation Ecosystem of Social Messaging Platforms* (NATO Strategic Communications Centre of Excellence, April 2020).

22. Numbers are estimates based on open source information collected during the spring of 2020.
23. For other categorizations, see for example: C. Foreman, '10 Types of Social Media and How Each Can Benefit Your Business', *Hootsuite*, 20 June 2017. [Retrieved 8 October 2020]; M. Storm, '5 Types of Social Media and Examples of Each', *WebFX*, 1 April 2020. [Retrieved 8 October 2020]; 'The 7 Different Types Of Social Media', *Biteable*, 21 March 2018. [Retrieved 8 October 2020].
24. Pamment et al., *Countering Information Influence Activities*.
25. C. Michel, 'How the Russians pretended to be Texans—and Texans believed them', *The Washington Post,* 17 October 2017. [Retrieved 8 October 2020].
26. Ibid.
27. M. N. Hussain, S. Tokdemir, N. Agarwal & S. Al-Khateeb, 'Analyzing Disinformation and Crowd Manipulation Tactics on YouTube', *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM),* 1092–95;
28. K. Conger, 'YouTube Disables 210 Channels That Spread Disinformation About Hong Kong Protests', *The New York Times*, 22 August 2019. [Retrieved 8 October 2020].
29. J. Li, 'China's messaging against the Hong Kong protests has found a new outlet: Pornhub', *Quartz*, 13 November 2019. [Retrieved 8 October 2020].
30. Manokara & Paramonova, *Manipulation Ecosystem of Social Messaging Platforms*.
31. M. Al Darwish, (2019, September 11). 'From Telegram to Twitter: The Lifecycle of Daesh Propaganda Material', VOX Pol, 11 September 2019. [Retrieved 8 October 2020].
32. N. Agarwal, & K. K. Bandeli, 'Blogs, Fake News, and Information Activities' in *Digital Hydra: Security Implications of False Information Online* (NATO StratCom Centre of Excellence, 2017), p. 31–45.
33. A Potemkin village of evidence refers to a false institutional network controlled by an actor conducting information influence, which involves multiple sources of information, such as news websites, blogs and social media pages. Such networks serve to plant and legitimize false information. See Pamment et al., *Countering Information Influence Activities*.
34. Agarwal & Bandeli, *Digital Hydra*.
35. B. Decker, 'Mapping the Anti-5G Campaign', *Global Disinformation Index*, 31 May 2019. [Retrieved 4 June 2020].
36. Ibid.
37. K. Collier, 'Reddit says UK election document leak points to Russia', *CNN*, 7 December 2019. [Retrieved 8 October 2020].
38. L. Hautala, 'Reddit uncovers Russian campaign to spread leaked UK documents', *c|net*, 6 December 2019. [Retrieved 8 October 2020].
39. [u/worstnerd], 'Suspected Campaign from Russia on Reddit' [Online forum post], *Reddit*, 6 December 2019. [Retrieved 8 October 2020].
40. Manipulation Service Provider refers to a company or an actor who sells manipulation of digital platforms, such as fake likes, shares and comments. See S. Bay & R. Fredheim, *Falling Behind: How Social Media Companies are Failing to Combat Inauthentic Behaviour Online*, (NATO StratCom Centre of Excellence, 2019).
41. PA Media,' TripAdvisor is failing to stop fake hotel reviews, says Which?', *The Guardian*, 6 September 2019. [Retrieved 12 October 2020].
42. O. Butler, 'I Made My Shed the Top-Rated Restaurant on TripAdvisor' [Blog post], Vice, 6 December 2017. [Retrieved 8 October 2020]; C. Ngak, 'Then and Now: A history of social networking sites', CBSNews, 6 July 2011. [Retrieved 8 October 2020].
43. C. Ngak, 'Then and Now'.
44. Myspace.com Competitive Analysis, Marketing Mix and Traffic, alexa.com, 2020. [Retrieved 8 October 2020].
45. NATO StratCom COE recently published a case study of three South-east Asian countries where the vari-

ations in social media patterns are showcased. See: J. C. Ong & R. Tapsell, *Mitigating Disinformation in Southeast Asian Elections: Lessons from Indonesia, Philippines and Thailand* (NATO StratCom Centre of Excellence, 2020).

46. World Map of Social Networks, vincos.it, 2020. [Retrieved 8 October 2020].
47. J. Clement, Social media penetration worldwide 2020, statista.com, 14 February 2020. [Retrieved 8 October 2020].
48. World Map of Social Networks
49. I. Koiranen, T. Keipi, A. Koivula & P. Räsänen, 'Changing patterns of social media use? A population-level study of Finland', *Universal Access in the Information Society* 19 (2019): 603–17.
50. Ibid.
51. A. Viens, 'This graph tells us who's using social media the most', World Economic Forum, 2 October 2019. [Retrieved 8 October 2020].
52. Ibid.
53. 'Demographics of Social Media Users and Adoption in the United States', *Pew Research Center*, 12 June 2019. [Retrieved 8 October 2020];
54. A. Perrin & M. Anderson, 'Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018', *Pew Research Center*, 10 April 2019.. [Retrieved 8 October 2020].
55. 'YouTube by the Numbers: Stats, Demographics & Fun Facts', omnicoreagency.com, 2020 [Retrieved 8 October 2020].
56. 'Demographics of Social Media Users and Adoption in the United States'.
57. Perrin & Anderson, 'Share of U.S. adults using social media'.
58. N. Thompson, 'Goodbye Gab, a Haven for the Far Right', *Wired*, 29 October 2018. [Retrieved 8 October 2020].
59. M. Zuckerberg, 'A Privacy-Focused Vision for Social Networking', *Facebook Notes,* 6 March 2019. [Retrieved 8 October 2020].
60. Ibid.
61. Ibid.
62. T. Buchanan, 'Why do people spread disinformation on social media?' (CREST Centre for Research and Evidence on Security Threats, 4 September 2020).
63. Bay & Fredheim, *Falling Behind*.
64. R. S. Goldzweig, B. Lupion & M. Meyer-Resende, *Social Media Monitoring During Elections*, (Open Society Foundations, 2019). [Retrieved 8 October 2020].
65. Ibid.
66. Pamment et al., *Countering Information Influence Activities*.
67. J. Pamment, H. Twetman, A. Fjällhed, H. Nothhaft, H. Engelson & E. Rönngren, *RESIST Counter Disinformation Toolkit*, (UK Government Communication Service, 2019). [Retrieved 8 October 2020].
68. Ibid.
69. A. Bredava,' Social Media Monitoring 101: All You Need to Know' [blog post], awario.com, 6 August 2018. [Retrieved 8 October 2020]
70. B. Ramez, 'Social Network APIs: The Internet's Portal to the Real World', toptal.com, 1 August 2016. [Retrieved 8 October 2020].
71. Z. Kharazian, 'Walkthrough: Tracking the spread of the debunked Plandemic video', DFRLab, 28 June 2020. [Retrieved 8 October 2020].
72. See for example: Yi Chang et al., "Ups and Downs in Buzzes: Life Cycle Modeling for Temporal Pattern Discovery," in *2014 IEEE International Conference on Data Mining* (2014 IEEE International Conference on Data Mining (ICDM), Shenzhen, China: IEEE, 2014), 749–54; Andrey Chinnov et al., "An Overview of Topic Discovery in Twitter Communication through Social Media Analytics" (Twenty-first Americas Conference on Information Systems, Puerto Ric, 2015),; Takako Hashimoto et al., "Event Detection from Millions of Tweets Related to the Great East Japan Earthquake Using Feature Selection Technique," 2015.
73. Emilio Ferrara and Zeyao Yang, "Quantifying the Effect of Sentiment on Information Diffusion in Social

Media," *PeerJ Computer Science* 1 (June 19, 2015)..

74. Pamment et al., *Countering Information Influence Activities*

75. See for example: F. B. Keller, D. Schoch, S. Stier, J. Yang, 'Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign' *Political Communication,* Volume 37 Issue 2 (2019), 256–80.
76. *Report of the Select Committee on Intelligence, United States Senate on Russian Active Measures, Campaigns and Interference in the 2016 U.S. Election—Volume 1: Russian Efforts Against Election Infrastructure with Additional Views*, 116th Congress Senate Report, 2019. [Retrieved 8 October 2020].
77. L. Townsend & C. Wallace, 'Social Media Research: A Guide to Ethics', (Economic and Social Research Council and the University of Aberdeen, 2016). [Retrieved 8 October 2020].
78. D. Funke & D. Flamini, 'A guide to anti-misinformation actions around the world', Poynter, 2020. [Retrieved 8 October 2020].
79. See: 'Steps of the research process', and excerpt from D. Blankenship, *Applied Research and Evaluation methods in Recreation* (Human Kinetics, 2010). [Retrieved 12 October 2020];
80. C. Dotto & R. Smith, *Newsgathering and Monitoring on the Social Web* (First Draft, 2019). [Retrieved 8 October 2020].
81. Johnston and Johnston argue that the model is limited in its ability to recognise biases and cannot fully incorporate theories of change. For a more detailed discussion, see: J.M. Johnston & R. Johnston (2008, June 28). 'Chapter Four: Testing the Intelligence Cycle Through Systems Modelling and Simulation' in R. Johnston, *Analytic Culture in the U.S. Intelligence Community* (Washington DC: Center for the Study of Intelligence, Central Intelligence Agency, 2005). [Retrieved 8 October 2020].
82. M. Phythian, *Understanding the Intelligence Cycle* (London: Routledge, 2015).
83. Johnson and Johnston, 'Chapter Four: Testing the Intelligence Cycle Through Systems Modelling and Simulation'
84. A. Liska, 'What is intelligence?' in *Building an Intelligence-Led Security Program,* (Elsevier Inc., 2015), p. 21–38.
85. 'Operationalization—Defining Variables Into Measurable Factors', explorable.com. [Retrieved 28 May 2020].
86. 'The Intelligence Cycle', fas.com, Federation of American Scientists, n.d. [Retrieved 8 October 2020].
87. *Guide for Civil Society on Monitoring Social Media During Elections* (Democracy Reporting International, European Commission, 2019). [Retrieved 8 October 2020].
88. 'The Intelligence Cycle'.
89. Pamment et al., *RESIST Counter Disinformation Toolkit*.
90. 'The Intelligence Cycle'.
91. M. Grandjean, M. (2016). 'A social network analysis of Twitter: Mapping the digital humanities community', *Cogent Arts & Humanities Volume* 3 Issue 1 (2016); S. Guarino, N. Trino, S. Chessa & G. Riotta, 'Beyond Fact-Checking: Network Analysis Tools for Monitoring Disinformation in Social Media' in *Complex Networks and Their Applications VIII,* Studies in Computational Intelligence series Volume 881 (2019): 436–47; E. Otte & R. Rousseau, 'Social network analysis: A powerful strategy, also for the information sciences', *Journal of Information Science,* Volume 28 Issue 6 (2002); 441–53;
92. B. Bhutani, N. Rastogi, P. Sehgal & A. Purwar, 'Fake News Detection Using Sentiment Analysis' *2019 Twelfth International Conference on Contemporary Computing (IC3)*.
93. P. Iosifidis & N. Nicoli, 'The battle to end fake news: A qualitative content analysis of Facebook announcements on how it combats disinformation', *International Communication Gazette,* Volume 82, Issue 1 (2019): 60–81.
94. J. Villena, 'Text Analytics to Detect Fake News', *MeaningCloud*, 16 October 2019. [Retrieved 8 October 2020].
95. J. Tompkins, 'Disinformation Detection: A review of linguistic feature selection and classification models in news veracity assessments', *arXiv preprint: 1910.12073*, 26 October 2019. [Retrieved 8 October 2020]; G. A. Mack, S. G. Eick & M. A. Clark, 'Models of Trust and Disinformation in the Open Press from Model-Driven

Linguistic Pattern Analysis' *2007 IEEE Aerospace Conference,* p. 1–12; F. T. Asr, 'One potential route to flagging fake news at scale: Linguistic analysis', NiemanLab, 16 August 2019. [Retrieved 8 October 2020].

96. E. Ferrara, 'Disinformation and social bot operations in the run up to the 2017 French presidential election', *First Monday,* Volume 22, Number 8 (2017).
97. 'How Geo-Location Analytics Is Being Used In Marketing', *Principa*, 23 March 2018. [Retrieved 8 October 2020].
98. K. Sharma, S. Seo, C. Meng, S. Rambhatla & Y. Liu, 'Covid-19 on social media: Analyzing misinformation in Twitter conversations', *arXiv preprint: 2003.12309*. 26 March 2020. [Retrieved 8 October 2020].
99. R. S. Nickerson, 'Confirmation bias: A Ubiquitous Phenomenon in Many Guises', *Review of General Psychology* Volume *2*, Issue 2 (1998): 175–220.
100. 'Social Media as a Research Tool—The Danger of Selection Bias', *IQS Research*, n.d. [Retrieved 8 October 2020].
101. H. Chen, Z. E. Zheng & Y. Ceran, 'De-Biasing the Reporting Bias in Social Media Analytics', *Production and Operations Management,* Volume 25, Issue 5 (2015): 849–65.
102. F. Morstatter & H. Liu, 'Discovering, assessing, and mitigating data bias in social media', *Online Social Networks and Media,* Volume 1 (2017): 1–13.
103. P. Krishnamurthy, 'Understanding Data Bias', *Medium*, 12 September 2019. [Retrieved 8 October 2020].
104. Pamment et al., *RESIST Counter Disinformation Toolkit*.
105. *CIA, A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis*, US Government, 2009.
106. R. J. Heuer Jr, R. j. Heuer, & R. H. Pherson, *Structured Analytic Techniques for Intelligence Analysis* (Washington, DC: CQ Press, 2010).
107. Pamment et al., *RESIST Counter Disinformation Toolkit*.
108. The prices for the tools and services referred to here, vary depending on what parts of the service or tool is purchased and most services offer varying price plans. The price levels indicated here are based on a combination of openly available information about price ranges and consultations with practitioners who use these services and offer an indication of the price level of the service generally. For accurate quotes, check the website of the company you are interested in or reach out to them directly.
109. See: G. Palmer, 'A Defining Moment: Crimson Hexagon is Joining Brandwatch', *Brandwatch*, 4 October 2018. [Retrieved 8 October 2020].
110. R. Alvarez, 'The Next Decade of Social Media and Customer Experience Management: Q&A with VP of Marketing Yoli Chisholm of Sprinklr', sprinklr.com, 20 January 2020. [Retrieved 8 October 2020.

Prepared and published by the
NATO STRATEGIC COMMUNICATIONS
CENTRE OF EXCELLENCE

The NATO Strategic Communications Centre of Excellence (NATO StratCom COE) is a NATO accredited multi-national organisation that conducts research, publishes studies, and provides strategic communications training for government and military personnel. Our mission is to make a positive contribution to Alliance's understanding of strategic communications and to facilitate accurate, appropriate, and timely communication among its members as objectives and roles emerge and evolve in the rapidly changing information environment.

Operating since 2014, we have carried out significant research enhancing NATO nations' situational awareness of the information environment and have contributed to exercises and trainings with subject matter expertise.

www.stratcomcoe.org | @stratcomcoe | info@stratcomcoe.org