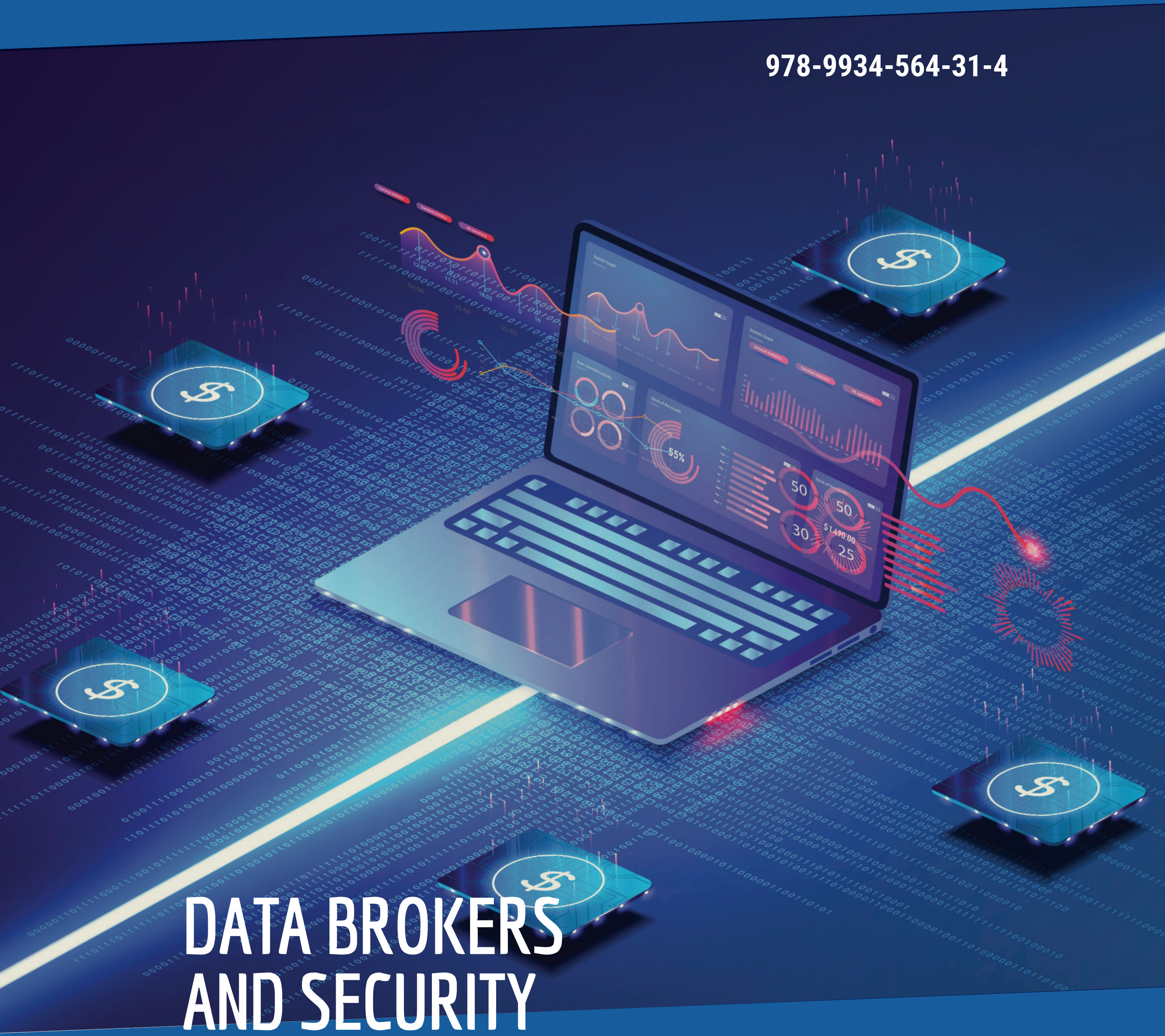
978-9934-564-31-4



DATA BROKERS AND SECURITY

Risks and vulnerabilities related to commercially available data

PUBLISHED BY THE
NATO STRATEGIC COMMUNICATIONS
CENTRE OF EXCELLENCE

ISBN: 978-9934-564-31-4
Authors: Henrik Twetman, Gundars Bergmanis-Korats
Project Manager: Henrik Twetman
Research: Singularex
Contributors: Nora Biteniece, Rolf Fredheim, Giorgio Bertolini, Sebastian Bay
Copy Editor: Anna Reynolds
Design: Jānis Uplejs

NATO StratCom COE does not endorse any of the products mentioned in the report.
This publication does not represent the opinions or policies of NATO or NATO StratCom COE.

PAGE LEFT BLANK INTENTIONALLY

Table of Contents

"You are not anonymous on this planet at this point in our existence. Everyone is trackable, traceable, discoverable to some degree."

Senior US Defense Department Official, *New York Times*, 20 December 2019[1]

INTRODUCTION

As connected individuals, part of a modern society, we generate data with everything we do. Every card-payment, website visit, browser search, social media post, and online message yields data points. Our phones register every action taken in every downloaded app; if the GPS is active they register every place we visit; and if we use biometric data monitoring they register our every heartbeat. It is virtually impossible to get an overview of the data we use and generate—data are everywhere.

The overwhelming abundance of data has ushered in the 'age of analytics',[2] where data informs the decisions, strategies, and activities of governments, corporations, and individuals.

For military organisations, data are a great asset— they provide valuable intelligence for operational planning, allow for near real-time situational awareness in the information environment, improve accuracy in recruitment, enable accurate simulations and exercises, and contribute to shortening military decision-making cycles.[3]

As with any new technology, data analytics create opportunities for both use and abuse. A number of risks and vulnerabilities have for too long been neglected in relation to the generation, collection, and dissemination of data.[4] When aggregated at scale, data reveal patterns and enable inferences that can compromise the integrity and threaten the security of individuals and organisations.

When traded without oversight, data can easily be used for unethical purposes. And when maliciously exploited, data can be used for tracking, manipulation, extortion, and scamming.[5] From a security perspective, the actors who control and own data in the information environment are critically important.

In previous reports, the NATO Strategic Communication Centre of Excellence (StratCom COE) has presented the issue of malicious use of data and digital information from a military perspective, showcasing risks mainly related to open source intelligence online and on social media.[6] These studies have focused on the direct exploitation of the digital domain by malicious actors, where these actors collect and process the data themselves. This report looks at another type of actor who fulfils a similar function but for commercial purposes—the data broker.

The Current Digital Arena and its Risks to Serving Military Personnel (2019)

Researchers studied the user data available in the digital environment and experimented with possible ways a malicious actor could exploit this data in the context of a military exercise. Their results suggest that an adversary would be able to collect enough personal data on soldiers to devise messages with precision, successfully influencing their chosen targets to carry out desired behaviours.

Camouflage for the Digital Domain: A Force Protection Framework for Armed Forces (2020)

In a previous report developed in collaboration with NATO Cooperative Cyber Defence Center of Excellence (CCDCOE) in Tallinn, we examined the broader risks of the digital domain and developed a framework for integrating digital risks into the military force protection framework. The conclusions we draw about data brokers in this report can be used together with this framework to identify, assess, and manage risks related to data brokers and the industry of data.

Data brokers are at the nexus of the data industry. They are, simply put, companies that collect, aggregate, and trade data for commercial gain. They do this somewhat covertly, taking advantage of the fact that users of online apps, media, and other platforms are largely unaware of how their personal data is collected, used, and sold in a relatively unregulated digital space with little transparency. They own and store billions of data points pertaining to anyone who inhabits the digital space, and then use these points to generate inferential data, placing the brokers among the most powerful organisations in today's digital world. As such, they are prime targets for exploitation by malicious actors, and their operations constitute a security risk.

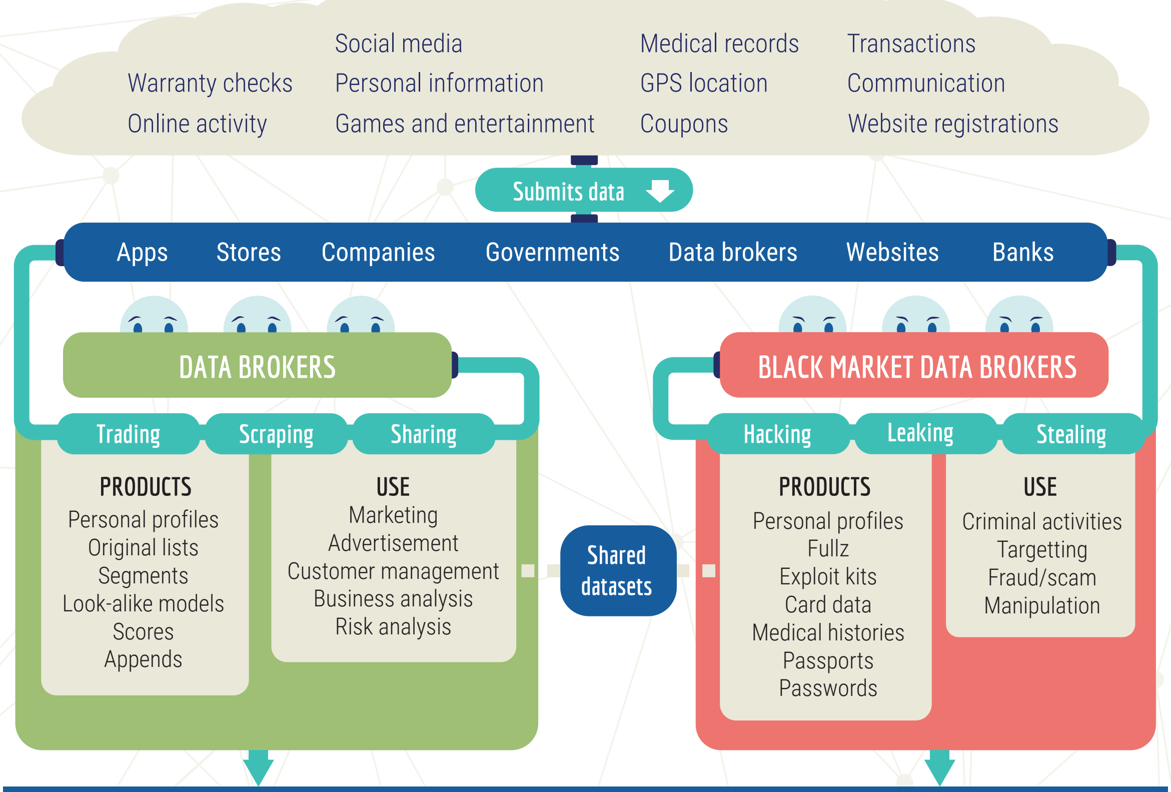
This report takes a closer look at data brokers and the data industry to investigate how the commercial availability of data can be exploited and lead to security issues for military organisations such as NATO and its Allies. It aims both to provide an overview of the data broker industry and its procedures, and to discuss risks and vulnerabilities related to this industry. It also describes the proof-of-concept experiment conducted by researchers from the NATO StratCom COE who engaged with multiple data brokers and purchased consumer data from an analytics company, and then used red-team analysis to assess how such data can be exploited.[7]

DATA BROKERS

"Companies whose primary business is collecting personal information about consumers from a variety of sources and aggregating, analysing, and sharing that information, or information derived from it, for purposes such as marketing products, verifying an individual's identity, or detecting fraud."

Federal Trade Commission, 2014[8]

DATA BROKERS AND THE INDUSTRY OF DATA



Warranty checks
Online activity
Social media
Personal information
Games and entertainment
Medical records
GPS location
Coupons
Transactions
Communication
Website registrations

Submits data

Apps · Stores · Companies · Governments · Data brokers · Websites · Banks

DATA BROKERS

Trading · Scraping · Sharing

PRODUCTS
Personal profiles
Original lists
Segments
Look-alike models
Scores
Appends

USE
Marketing
Advertisement
Customer management
Business analysis
Risk analysis

Shared datasets

BLACK MARKET DATA BROKERS

Hacking · Leaking · Stealing

PRODUCTS
Personal profiles
Fullz
Exploit kits
Card data
Medical histories
Passports
Passwords

USE
Criminal activities
Targetting
Fraud/scam
Manipulation

RISKS AND VULNERABILITIES

General issues and risks

Violation of Privacy
Data collected and used without consent by user
Potentially harmful inferential data
No transparency with regards to what data is kept on a user and for what reason
No control over who obtains personal data

Advertisement and manipulation
Unwanted advertisement
Political manipulation via targeted ads
Misuse of personal data

Data exposure
Lacking security of the data broker
Exposure of personal information
Criminal activities

Explotation
Spam
Phishing attacks
Identity theft
Scams

Potential risk to military organisations

Personnel
Extortion / black mail / dioxing
Manipulation of behaviour or opinion
Impersonation or identity theft
Intelligence gathering / surveillance

Equipment
Exposure of device IDs
Mapping of communication patterns
Credit card theft
Intelligence on equipment

Activity
Mapping or tracking of personnel movement
Geo-localisation of ops or exercises
Disruption of activities

Information
Information theft via for example spear phishing
Exposure of sensitive information such as lists of personnel etc.
Data leaks / hacks

Facilities
Geo-localisation of sensitive or secret facilities
Identification of personnell working in specific facilities
Access to facilities via impersonation

THE BUSINESS OF DATA BROKERS

In recent years, the development of new technologies and business models, such as social media and mobile applications, has increased the availability, variety, and volume of data. It has also led to a growth in businesses dealing in data. According to Gartner, an estimated 5,000 data brokers work worldwide. Some of the top players in this market are Acxiom, Experian, Equifax, CoreLogic, TransUnion, Oracle, Lifelock, H.I.G. Capital, PeekYou, and TowerData.[9] Nearly 10 million open datasets have been published by government agencies and non-governmental organisations (NGOs).[10]

> *"On a daily basis, consumers engage in a variety of online and offline activities that reveal personal information about them. [...] The entities they interact with collect information about them and, in many instances, provide or sell that information to data brokers."*
>
> *Federal Trade Commission, 2014*[11]

The primary business of a data broker is to archive personal information and to collect behavioural data from a variety of online and offline sources without interacting directly with users. The collected data is then aggregated and analysed for further use, for example for user profiling, which provides end-user institutions with a great deal of information about the lives of their consumers and has thus entirely restructured how enterprises do their business.[12]

The data sold by data brokers is not only useful and valuable, but immensely powerful. The ability to access such comprehensive data has transformational potential for businesses, governments, militaries, and even private individuals; in today's world, data is the basis of productivity, competition, and innovation.[13] In a report on the uses of big data, McKinsey lists five ways in which data create value: increasing transparency of information, more accurate analytics, tailoring of products and services, improved decision-making, and higher rates of innovation.[14] According to Oxylabs, data-driven organizations are up to twenty-three times more likely to acquire customers, six times more likely to retain customers, and nineteen times more likely to be profitable.[15]

The data industry in numbers[16]

- 4,8 billion internet users globally
- 175 zettabytes of data produced worldwide by 2025[17]
- 10 million open datasets
- 5 000 data brokers worldwide
- $178 billion—current revenues for the data broker industry
- $400 billion—projected worth of the global data economy by 2025

TYPES OF DATA

Data brokers have access to a myriad of data, most of which has been collected freely from public records or scraped online. But data brokers also purchase data from other commercial actors, such as social media platforms and apps, to complete their data sets, and collaborate with each other to construct custom data sets.[18] Some brokers have data-sharing partnerships with primary data collectors, or have access to scraped data, although recent privacy scandals have led to some tightening of privacy settings on behalf of the primary data collectors.[19]

The Big Fish[20]

The US-based data broker Acxiom (recently rebranded to LiveRamp) is considered one of the world's leading data brokers and serves as a point of reference for the scale of the industry. Acxiom reportedly has over 20 000 servers for collecting and analysing data on over 700 million individuals worldwide. WebFX reports that the company has up to 3000 data points for every US consumer and up to 1500 data points per person in their database. Acxiom themselves claim that the company "has the most expansive and compliant data offering in the world, which now encompasses more than 62 countries, 2.5 billion addressable consumers and more than 10,000 attributes—for a comprehensive representation of 68 percent of the world's online population".[21]

Not all data brokers collect the same data—they tend to specialize in certain industries or niche markets to gain a competitive advantage. For this reason, it is also common for brokers to trade data with each other in order to provide the most accurate possible data sets to their clients.

A broker's sources fall into three broad categories: publicly available sources, commercial sources, and online tracking data. In order to build more complete data points of a target subject, companies can combine data from multiple sources to create detailed profiles.[22]

Source	Examples of data
Public sources	Demographic information, property records, court filings, criminal convictions, professional licenses, census data, birth certificates, marriage licenses, divorce records, state professional and recreational license records, voter registration information, bankruptcy records, business listings, telephone books, vehicle registration records, etc.
Commercial sources	Purchase history, warranty registration, credit information, employment registration, loyalty card data, membership data, subscriptions, etc.
Online tracking	Social media profiles, web browsing activity, mobile apps, media reports, websites, mail-in rebate forms, forum posts, web browser cookies, plugins, addons, device data, IP fingerprints, network data, metadata, etc.

Brokers' data sets consist of two types of data: observed and inferred (or modelled). Observed data is data that has been collected and is actual, whereas inferred data is gleaned from observed data by modelling or profiling.[23] For example, if you have visited web pages related to bungy jumping, a broker might infer that you are a 'risk taker' with an interest in bungy jumping. In practice, this means observed data collected on an individual can be used to infer a great deal more data. Brokers typically collect not only what they immediately need or can use, but hoover up as much information as possible to compile comprehensive data sets that might have some future use.[24]

Building a digital profile—how it could work[25]

In a detailed report commissioned by the Norwegian Consumer Council [*Forbrukerrådet*], researchers describe in detail how data brokers construct digital profiles of individuals. Their description of how different personal identifiers are used to construct behavioural profiles spanning multiple devices for the purpose of targeting advertisement illustrates how data brokers construct their data sets:

"A data broker [...] may receive information about [an app] user, including his Android Advertising ID and his IP address. The user then opens [the app] while his phone is connected to his home Wi-Fi network. When this happens, the data broker can use the IP address of the home network to identify the user's home, and append this to the unique profile it is compiling about the user. If the user has a computer connected to the same network, this computer will have the same IP address. The data broker can then use the IP address to connect the computer to the same user, and identify that user when their IP address makes requests on other publisher pages within their ad network. Now the data broker knows that the same individual is using both the phone and the computer, which allows it to track behaviour across devices and target the user and their devices with ads on different networks."

DATA BROKER PRODUCTS

Data brokers aggregate data into marketable products, for example lists based on observed data, consumer segments based on a combination of observed and inferred data, and look-alike models used to predict the behaviour of a subject based on the past behaviour of similar subjects. Data brokers also create various scores (such as credit scores) that can be used to make predictions about the likely behaviour of existing or prospective customers.[26]

For example, a broker may include a subject with the data point 'interested in bungy jumping" in a group of other subjects interested in bungy jumping. As the Federal Trade Commission highlights, such 'segment data' can then be use with varying effect. If it is purchased by a travel agency, it can be used to offer that subject targeted advertisements. If it is purchased by an insurance company, however, the same data may be used to increase the insurance fee for that subject because he or she is considered to be more likely to engage in unsafe activities.

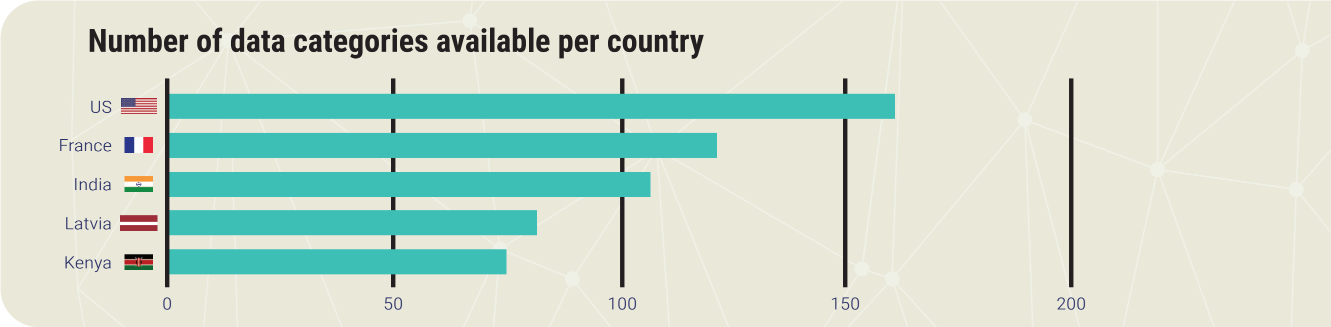
Data brokers tailor to many different customers and sectors. The largest markets for brokered data include advertisement and marketing, health education, credit and insurance, customer services, and government and law enforcement. These markets use data broker services for a variety of ends, but generally to improve their businesses or decision-making.

Case: Secret Service bypasses warrants using data brokers[27]

It is not only malicious actors that can exploit the services of data brokers. ArsTechnica recently reported on how law enforcement agencies in the US, including the Secret Service, use data brokers to purchase location data for which they would otherwise need a warrant. Outdated regulatory frameworks place no limitations on public agencies seeking to purchase data that would otherwise require a warrant, and there are few limitations on what brokers can collect, store, and sell, effectively short circuiting civil rights.

AVAILABILITY AND QUALITY

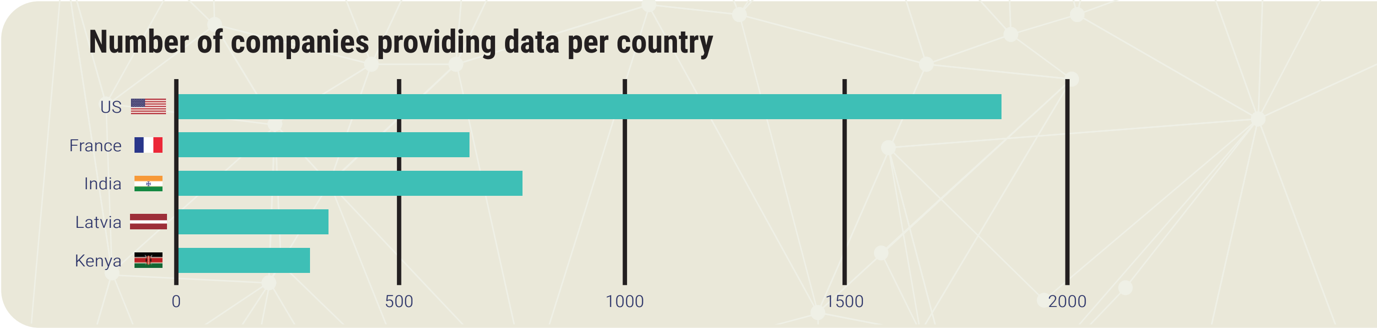
Despite the overwhelming abundance of data available generally, the data traded by brokers is not all-encompassing, even at the level of industry. Researchers used the search service provided by Datarade[28] to sample the availability of different data categories for five countries—the US, France, India, Latvia, and Kenya—to illustrate the variations within the industry.


Number of data categories available per country

Although India has the second largest population of Internet users in the world, data vendors operate with fewer data categories in India than in France or the US. Similarly, Kenya has more Internet users than Latvia, but has fewer data categories available about its citizens. This suggests it is not market size that determines the abundance of data vendors, but other factors such as demand and legislative landscape.

Looking specifically at location data, for example, we can see that there are several types, e.g. raw location data limited to latitude and longitude information, location data that includes certain calculated fields such as mobile state, speed, alleviation etc., and enhanced location data that includes information such as device make, device price, carrier, age range, gender, interest data and more. These three types are available in all five sampled countries to some extent. However, some data categories are available for only one or two countries, but not the others. These categories include, but are not limited to, psychographic data, which is available for the US and France, and court, litigation, life stage, and patient data, which are available for the US. The abundance of data available about US citizens can be at least partially explained by the lack of regulation to protect citizens' privacy. The aforementioned data categories are deemed to be sensitive in Europe.

When the total number of data providers within each data category are compared across countries, a similar picture appears. The US has the most vendors by a large margin.


Number of companies providing data per country

The data that brokers possess is not always accurate or up to date. This is not necessarily a problem for brokers or their clients, since the products they sell do not need to be 100% accurate to achieve clients' goals, for example reaching a specific consumer segment with marketing information. In a study of a range of US-based data brokers, the Federal Trade Commission found that brokers do not typically "assess [the quality of data from] government and other publicly available sources, [but] may take some steps to assess their commercial sources in order to ensure that the sources provide accurate data."[29] Our research indicates that quantity overshadows quality in the data broker industry, and that on average only 50–60% of data can be considered precise.

PURCHASING DATA FROM A DATA BROKER

The process of buying data depends on what type of data you are looking for and from which company you buy it. You can buy pre-packaged marketing data sets directly from a company webpage and get the data delivered to you via email in a matter of minutes. Higher-end service services have more complex procurement processes and their products are more commonly custom made.

Case: Buying your own data[30]

Journalist Caitlyn Renee Miller reported on her experience of purchasing online data about herself from a data broker in *The Atlantic*. She describes how she uploaded her email address and paid a $50 fee in exchange for "net worth, age, zip code, and education, among other personal information". A few hours later she received a report in her mailbox with a summary of the data associated with her email. To her disappointment, about 50% of the data she received was inaccurate. Caitlyn's experience illustrates the ease at which data can be purchased, but also reveals the quality-to-price ratio—cheaper services may not provide quality data.

Higher-end services will also screen their clients as part of the purchasing process. Screening may include general verification procedures, research, and face-to-face meetings.[31] Screening procedures are intended to confirm the legitimacy of the client and to guarantee lawful use of the product. However, this is not industry standard, and many brokers provide data without significant screening.

The services of higher-end data brokers are often regulated by contracts that define the parameters of the transaction in detail, whereas lower-end services are more easily accessible and less controlled.[32] Still, enforcing contractual obligations, such as restrictions on reselling data, is difficult and data brokers are limited in their will and/or ability to monitor how clients are using the data. In an interview with *Wired*, a representative of Oxydata is quoted as saying "there is no way for us to enforce all of our clients to follow the best data protection practices and guidelines".[33]

Pricing varies greatly depending on service. Simpler services have fixed prices for specific products, while the clients of higher-end services request quotes. A UK-based date brokerage describes its pricing model as follows:

> *There's no flat fee for data by volume, every brief is managed individually to get you the best deal for your budget. This entails shopping around to each supplier of data, sometimes using a combination of several, who all have different pricing brackets and data volumes.*[34]

We found that it is common for brokers to have a fixed base price for bespoke orders, with additional costs for specific requests. For example, when sampeled Onemata, a US-based data reseller, had a starting price of $20,000, whereas Tamoco, a UK-based location data reseller, had a starting price of £2500. The actual price of a dataset is based on several factors, including the number of months of historical data requested, the country or region covered, and the data attributes required. However, it is not always clear how these factors play into calculating the final price. The exclusivity of the data set is likely to have the greatest impact on price. An expert we consulted who had purchased a data set with historical data on improvised explosive device (IED) locations in conflict zones had been billed $200 000.

THE BLACK MARKET FOR DATA

While not fully transparent, most data brokers operate within the spectrum of legality. There is, however, also a black market for data where a different type of broker operates covertly and illegally on the dark web. The actual data traded without oversight on the black market may be the same data traded by legitimate brokers, but may also be unlawfully obtained data or more sensitive data sets that are not otherwise available.

Personal information, digital user profiles, log-in details, leaked records and databases, credit card information, medical records, and passports are some examples of data that can be bought on the black market. More advanced services include data bundles that combine different sets of data into a specific product designed for exploitation. The black market product slate is diverse, with a constantly increasing variety of goods and services.

Sellers on the black market are not always hackers. Data traffickers, market analysts, and technical experts all offer specialized services tailored to the buyer's needs.[35] In this sense, the black market for data is mature and professional—some actors even guarantee services and products to build trust with their customers and regularly conduct their own compliance checks.

Get the fullz[36]

In the 2019 report, *A Look Inside the Dark Web*, the cybersecurity company Armor sheds light on the black market data product known as a 'fullz'. A fullz is a bundle of data that contains "the motherload [*sic.*] of personal information on an individual". For a bit of extra money, data such as credit card numbers, bank accounts, and security questions and answers can be purchased. In essence, a fullz provides the buyer with everything needed for criminal activities such as scamming, fraud, impersonation, or identity theft. The price of a fullz varies between $17 and $60. The Armor report is an excellent resource for information on services available on the dark web.

The black market for data is part of the larger online black market where hackers and cyber criminals trade goods and services. Like any other market, it has a hierarchy with a variety of actors and different levels of access. Low-end services make up the majority of the black market for data;[37] these are easily accessible and open to anyone. Brokers can be found directly via search engines such as Google and purchases are made using a credit card or online payment service. Brokers dealing in high-end services are more selective and are often accessible by invitation only or after careful vetting.[38]

The covert nature of the black market makes it difficult to determine its size and structure, but available evidence indicates that the market is huge and growing. A study by Bromium Inc. estimates the revenue of the entire cybercrime market, of which the black market for data makes up a sizable proportion, as $1.5 trillion.[39] It is possible that the black market for data is more profitable than the entire legitimate data industry. The scope of the black market is further illustrated by the number of intrusions and breaches. According to the 2019 Year End Data Breach QuickView Report, over 7,000 data breaches were reported and over 15.1 billion records exposed in 2019 alone.[40] Similarly, Digital Shadows found that in 2019 roughly 150 million sensitive records had been exposed online, the highest number being in the banking sector, closely followed by the medical/healthcare sector.[41] These records are actively traded on the black market.

Case: The price of your data on the black market[42]

In a blogpost on Experian, Brian Stack took inventory of the prices at which data are traded on the dark web. He found that US social security numbers can be purchased for as little as $1, while medical records and passports cost up to $2000. Other types of data available through these services include online payment history and login info, credit card details, and drivers' licenses and diplomas. Such data can be bought either as a one-off or in bulk.

Illustration: Comparing the white and the black markets for data

Characteristic	White market	Black market
Specialization	In most cases, traded data is used to understand and subsequently influence a target group.	As-a-service models unique to each customer; customized malware. Uses are likely to be illegal.
Market size by revenue	Approximately $200 billion	Approximately $1.5 trillion
Reliability	Both markets engage in questionable practices with regard to legality and service. Sellers' reputation are vital for the business.	
Government regulation and impact	Government in constant confrontation with brokers, arguing that they operate on the verge of law.	Government in an arms race with cybercriminals, unable to curb the continual rapid expansion of the black market.
Sophistication	Technological innovation is a key factor in market development, the goal of which is to collect and effectively analyse as much information as possible.	Continuous market-driven improvements in security, anonymity, and more sophisticated ways to please customers.
Competition	Highly competitive market dominated by established strong players.	Easy to start a business in the ever-expanding black market; success depends on reputation and skills.

> Location data, like any other data set, has the possibility to be misused by bad actors, but has perfectly legitimate and ethically positive use cases which unfortunately go unreported by the media.
>
> Tamoco [a data broker], *NRK*, 3 June 2020[43]

DATA AND SECURITY

While the benefits of data should not be understated, the risks related to data collection, especially the commercial availability of data, should not be overlooked. Recent scandals have triggered a broader discussion about the security risks associated with the commercial availability of data and how the industry of data can be regulated.[44] The following sections explore some of the security issues related to data brokers, before turning to risks related to the malicious exploitation of military data.

GENERAL SECURITY ISSUES WITH DATA BROKERS

As we have discussed, multiple security-related issues arise from the commercial trade of data generally, and the practices of data brokers specifically. Here, we highlight some of the most common security risks related to personal integrity, data leaks, intrusive advertising, and exploitation on the levels of the individual and of society. Understanding these risks and harms is vital for creating an appropriate regulatory framework for the data collection industry;[45] this discussion provides a baseline for understanding the risks faced by governments and militaries risks considered in the following section.

> “While everyone is looking at cybercriminals and the Dark Web, Data Brokers are doing far more damage to people’s privacy in plain sight.”
>
> Daniel Miessler, ‘The Dark Web has Nothing on Data Brokers’, 25 June 2020.[46]

1. Violation of Privacy

Data brokers collect, process, and trade data about private individuals without their explicit consent, and often without their knowledge or awareness. Consumers sometimes realise they are submitting data that can be used by third-party vendors, but it tends to be difficult for a user to fully understand what they are agreeing to; consent to share data is often "sneakily included".[47] Industry practices enable unexpected transfers of data—a user may consent to submit personal data to a company (for example by accepting a cookie on a website) without realizing that years later, when that company has changed owners or gone bankrupt, the same data points may belong to a different company with completely different business practises.

Data brokers sometimes dismiss these practises by referring to the anonymous nature of the data they store and sell but, as has been demonstrated many times, anonymous data can easily be de-anonymized by cross-referencing different data sets.[48] A study published in *Nature* showed, for example, how under the right conditions geolocation data points are sufficient to identify 95% of individuals in a data set.[49]

Data brokers constitute a direct threat to privacy that can lead to the exposure of sensitive personal information or to a situation in which an individual may be harmed without knowing why and without the possibility of rectifying the issue,[50] for example being denied a mortgage, given a higher insurance rate, or being the subject of discriminatory practices.

Because of the lamentable lack of transparency and regulation, there is no way of knowing what information brokers have about you and who this information may be sold to. Testifying before congress, Pam Dixon, the executive director of the World Privacy Forum, described modern data brokers as lacking restraint in that they provide products such as lists of rape victims, those who use domestic violence shelters, and sufferers from genetic disease, which clearly violates the integrity of personal information.[51]

Case: Daughter killed in car crash[52]

The aggregation and trade of disparate data can sometimes cause problems related to integrity. An example of this was a promotional letter that OfficeMax sent to a customer in 2014. The second line included the words "Daughter killed in car crash", causing the recipient to feel revictimized over a year after his daughter had passed away. OfficeMax explained that the incident as an "inadvertent error" caused by a "mailing list rented through a third-party provider". The incident also made the customer question how much data OfficeMax and the data broker had about him, and in what context his daughter's passing had been registered in the database and was deemed relevant for business.

2. Data exposure

While trading and selling of personal data often violates the privacy of individuals, the exposure of sensitive information by a hack or leak can be even more damaging. Because they store vast amounts of valuable information, data brokers become prime targets for hackers and cybercriminals. There have been multiple high-profile cases recently that illustrate this problem, in which personal data was extracted from data brokers, largely due to their insufficient cybersecurity practices.[53] Almost every top broker has been hacked at some point: Acxiom was hacked in 2003, Epsilon in 2011, and Experian in 2015, to name just a few of the cases that have reached the public's awareness.[54] The lack of security practices is putting billions of peoples' private and sensitive data at risk.

Case: Data broker exposes millions of online profiles[55]

The tech-journal *Comparitech* recently reported on a data brokerage that accidentally exposed its database of 235 million social media profiles, which included personal information and contact information that was uploaded without proper security. Much of the data originated from the now defunct company Deep Social and had been scraped from social media platforms such as YouTube, TikTok, and Instagram, seemingly at odds with the terms of use of these platforms. *Comparitech* reported that "even though the information is publicly available, the size and scope of an aggregated database makes it more vulnerable to mass attack than it would be in isolation", and that the data contained in the exposed database could be used for spam, phishing, impersonation, scams, and even for illegitimate facial recognition.

3. Advertisement and manipulation

Brokers sell data mainly for marketing purposes and for targeted advertisement. While this may be harmless and even useful in many cases, it can also cause problems for individuals who are subjected to manipulative marketing practices for unwanted products and services or political manipulation based on their personal data. In a report by the Norwegian Consumer Council the practices of the ad tech industry (within which data brokers play a central role) are described as "out of control, rife with privacy violations and breaches of European law, and highly problematic from an ethical perspective".[56]

Case: Cambridge Analytica[57]

The Cambridge Analytica scandal of 2016 demonstrated how personal data collected for advertising purposes can be blatantly misused for purposes of political manipulation. The political data analysis company used improperly obtained social media data for over 87 million individuals to construct psychological profiles used for targeted political advertisement. The company's services were used by both Donald Trump's presidential campaign and the Brexit Leave campaign to sway voters.

4. Exploitation

Finally, personal data can be used in harmful and unethical ways, such as to facilitate identity theft or for purposes of scamming. This is true of data from legitimate brokers as well as from their shadowy twins on the black market, although the threshold for exploitation is much lower for black market services. Even simpler products such as email-lists and addresses can be used for phishing attacks and scams, whereas more detailed personal profiles enable sophisticated exploitation such as identity theft or credit card skimming.

Case: The invitation-only black market for stolen data[58]

The case of the Genesis Store, now shut down, provides us with some insight into how the black market for data functions. According to the experts at Kaspersky Lab, this store used an invitation-based online platform to sell stolen user profiles, including browser fingerprints, cookies, logins with passwords, and credit card information. Prices ranged from $5 to $200 per profile. These were calculated automatically using an algorithm that would value each profile on the basis of a range of variables. The most interesting aspect of Genesis Store's service was their browser plugin that made it easy to install a stolen digital profile to generate a digital double of the source user.

> "If you can get information on someone online, you might be able to impersonate them or use their credit history, or perhaps get into a password-protected website if you can answer security questions about people."
>
> Paul Stephens, Director of Policy and Advocacy at Privacy Rights Clearinghouse[59]

THE MALICIOUS USE OF DATA

The issues related to the exploitation of personal data legally sold by brokers merit concern. These issues often arise from structural inadequacies within the industry and the lack of appropriate regulation; they lead to problems at a relatively small scale, the scale of the individual. More serious risks arise when data brokers are exploited by malicious actors, such as hostile states or terrorist groups, with the intention to cause harm. In such cases, data brokers constitute a weak link for national security.[60]

Data brokerages are a treasure trove for malicious actors in the 21st century, especially from a military perspective. Without costly intelligence and reconnaissance capacities, a malicious actor can obtain detailed and potentially sensitive information about its targets. Without concern for the legality of information collection, vast and detailed data sets can be obtained immediately and at a comparatively cheap price. And because the industry has very low barriers to entry and only sporadically conducts screenings, the market is open to any actor with the means to pay for products and services. If access cannot be obtained legally, hacking into a data broker's server is also lucrative, since a wealth of data is stored in one place and security practices tend to be insufficient.

This section explores potential risks associated with the malicious use of data from a military perspective.

Case: Hacked data broker accounts exploited[61]

As they hold the keys to accessing vast amounts of data stored on their servers, the data brokers themselves can become targets of malign activity. Amidst the Covid-19 pandemic, a little-known US data broker called Interactive Data LLC was targeted by a malicious actor who gathered personal data on people and business later used for impersonation, scams, and fraudulent emails. The potential for the exploitation of personal data is huge, and data brokers who routinely collect, aggregate, and store this type of data are naturally prime targets for malicious actors who can either purchase the data legally or obtain it illegally and then use it to achieve their objectives.

A TAXONOMY OF RISK FOR MILITARY ORGANISATIONS

From NATO's perspective, any internal or external factors that contribute to uncertainty with regard to whether and when Allied forces can achieve their objectives are considered risks.[62] Risk can be understood as "a function of threats exploiting vulnerabilities to obtain, damage, or destroy assets"[63] and it can have implications at all levels of operations.[64]

While private individuals will want to protect 'assets' such as private or sensitive data, a military organization considers certain categories of assets to be critical and in constant need of protection—these include personnel, equipment, information, facilities, and activities.[65] Threats and vulnerabilities to these assets can be identified at different levels, arising from data generated by military personnel, the structure of the datamarket, or from a specific data broker/data brokerage.

> "Commanders should identify, assess and manage the risks involved in their military operations, and provide guidance to the staff and subordinates for risk reduction, mitigation and exploitation."
>
> NATO AJP-3 Ed C v.1[66]

The risk taxonomy developed here extrapolates vulnerabilities and threats based on the five categories of military assets mentioned above; this enables military organisations to identify and assess risks. This taxonomy is by no means exhaustive. Vulnerabilities and threats vary over time and between different military organisations. There is not always a clear division between the vulnerabilities and threats associated with various assets. For example, the personal information highlighted under 'personnel' could be used to identify a specific military facility, and so on.

Still, the taxonomy is useful in that it shows how data obtainable via a broker can be used maliciously to cause harm or to damage military assets. Each category of assets is described below together with examples of how data can be exploited. To be sure, not all of the threats discussed below arise only from the bad practices of data brokers, but in all cases brokers constitute a weak link in the security chain.

Risk Taxonomy

Asset	Vulnerabilities	Threats
Personnel	- Personal information including social media data on personal preferences - Geolocation data	- Extortion/black mail / Doxing - Manipulation of behaviour or opinion - Impersonation or identity theft - Intelligence gathering / Surveillance
Equipment	- Device ID - Data on device use - Data on specific equipment such as credit cards	- Exposure of device IDs - Intelligence on equipment - Mapping of communication patterns - Credit card theft
Information	- Information about personnel - Information about system usage and user behaviour	- Information theft via for example spear phishing - Exposure of sensitive information such as lists of personnel etc. - Data leaks/hacks
Facilities	- Geolocation data - Personal information - User data	- Localisation of sensitive or secret facilities - Identification of personnel working in specific facilities - Access to facilities via impersonation
Activities	- Geolocation data - Personal information including social media data on personal preferences	- Mapping or tracking of personnel movement - Localisation of ops or exercises - Disruption of activities

Personnel

Personnel is a critical asset for any organisation. Risks related to personnel interlink with risks to other assets, and human error is often the weakest link in a security chain. Considering the personal nature of much of the data and services provided by data brokers, the risks to personnel are apparent. On a basic level, personal information can be used to identify personnel working with specific tasks within the military, or to track their movement and behaviour as part of hostile intelligence gathering—purchasing this information from a data broker is more cost effective than traditional spycraft. Sensitive data, such as financial information or private internet habits, can be used to blackmail or extort decision-makers. Behavioural data and information about personal preferences is useful for fine-tuning hostile influence activities and manipulation, and personal profiles can be exploited for purposes of impersonation or identity theft.

Case: Tracking individuals using anonymous data[67]

In their impressive article series on data integrity, the *New York Times* showed how easily anonymous geolocation data can be deanonymized and used to track individuals. Having obtained a dataset of some 50 billion location pings, journalists were able to identify, among others, individuals belonging to the President's Secret Service detail. They describe the process "as easy as combining home and work locations with public information". This clearly illustrates the security issues arising from the collection and aggregation of big data sets. The *New York Times* describes the industry of data as a "system in which data surveillance practices are hidden from consumers and in which much of the collection of information is done without the full knowledge of the device holders."[68]

Equipment

Military equipment includes any type of material or gear used within a military organisation, ranging from combat vehicles to mobile phones. From a risk perspective, equipment is vulnerable to threats such as intelligence gathering about capacities and capabilities, manipulation, and sabotage. Data available via brokers can threaten material assets in a variety of ways. For example, device ID data can be correlated with personal information and geolocation data to map user patterns and communication infrastructure. Purchase history and credit card details can be exploited both for intelligence purposes and for theft. Correlation of geolocation data with open-source information, such as social media posts and images, can further be used to assess the capacities and capabilities of a military unit, as the example below shows.

Case: Beer app used to track military personnel[69]

Using the beer-rating app Untappd, open-source investigators from Bellingcat managed to track the location history of military personnel who voluntarily, but most likely unknowingly, had been sharing sensitive data about their location and, in some cases, pictures of sensitive military equipment. Automatically generated location data can be cross referenced with uploaded photographs to confirm positions within military bases and map patterns of movement over time, illustrating how security issues related to data often develop in the nexus between digital technologies and human interaction. Often, users are not even aware of what data they submit and how easily it can be exploited.

Information

Securing restricted or sensitive information is critical for any military organisation, and there are multiple ways in which commercially available data can pose a threat to the integrity of information. For example, profile data and geolocation data can be cross referenced to identify key personnel, whose personal data profiles can be used to design and direct spear phishing attacks in order to gain access to restricted systems or obtain sensitive information. A simple data broker product such as an email list for known military personnel is useful for this purpose. The services offered on the black market for data, which are often tailored to scammers, make intrusions even easier. The aggregation of data by a broker may also expose sensitive information, such as which personnel work at a secret facility.

Facilities

As with personnel, geolocation data can easily be exploited to reveal the location of military facilities and information about their personnel and functions. In a well-publicised case, journalists used the fitness app Strava to locate military bases abroad by looking at GPS data for exercise routes.[70] Data brokers sell similar data sets from a variety of applications and sources, but with more data points. Personal data profiles can be used to impersonate those with access to restricted facilities. The example below is perhaps a sneak peak into the future of data-driven risks.

Case: Using images to steal fingerprints[71]

Data can sometimes be exploited in unexpected ways. Researchers from Japan have discovered that images including pictures of your fingers can be used to recreate fingerprints, which in turn can be used to bypass biometric security systems. They noted this threat was particularly big in Japan, where it is popular to upload 'V- sign' selfies that expose the fingertips to the camera. This is one of many ways in which online photos and photo databases can be exploited. Data from a data broker could be used to identify which personnel have access to a particular facility, so that their profiles can be scraped on social media for data useful for replicating biometrics.

Activities

Finally, the activities of military organisations can be disrupted using commercially available data. Geolocation data can be used to identify where military exercises or operations are being held, and tracking data can provide intelligence on the nature and scope of military activities. The case study below illustrates these risks. Personal data on military personnel can also be used to disrupt, for example, by directing influence activities against exercises or operations.

Case: Tracking military personnel using mobile data[72]

Norwegian journalists from *NRK* wanted to test if they could use mobile phone data to track military personnel, and bought a data set containing location data from 140 000 phones from a British broker for NOK 35 000 (roughly equivalent to $3 800). From this data they could easily see which phones are regularly located at military bases, and then follow how they were moving after work hours, to discern where personnel working at the bases were living. The broker that sold *NRK* the allegedly anonymous data is currently under review by both British and Norwegian data protection authorities after *NRK* published its report.[73]

EXPERIMENT

There are many proof-of-concept experiments showing how data brokers can be exploited and how easy it is to manipulate, deanonymize, and leverage data sets containing personal or sensitive information for malicious purposes. We have already given many examples, the most notable, perhaps, being the *New York Times* exposé on the ease of locating the President's security detail using anonymous geolocation data.[74] These cases highlight the vulnerabilities and threats inherent in the industry as it is structured today. However, these experiments have been mostly conducted in the US, using data categories that may be significantly harder to source in other countries.

We set out to test whether the results of these experiments are applicable in other contexts. Namely, we wanted to understand what data can be purchased in a smaller European country (in this case Latvia), how easy it is to purchase, what it would cost, and how the data would be delivered to us.

To this end, we reached out to several data brokers operating in Europe to sample their services. We collected information about data availability, quality, and pricing. We also purchased a marketing analysis of a specified target audience from an analytics company to see how this data could potentially be exploited using a red-team analysis model.[75] For this experiment, we did not purchase any raw data and the experiment does not provide an encompassing review of the market for data in Latvia, but it does provide some insight into how the marketplace works and what data can be obtained.

AVAILABILITY, QUALITY AND PRICE

While it has been well established that there is an abundance of data available from brokers generally, this does not always hold up for smaller markets such as Latvia. Using the Datarade platform,[76] a service that searches brokerages and the data they sell, we found a total of 82 data categories regarding Latvia sold by some 300 different companies (see Chapter 2 for graphs). The range of categories offered was comprehensive—including data on physical addresses, IP addresses, business listings, marketing, real estate, and pharma—but their usefulness was limited by the fact that many of the categories contained only a single data set, or a few of very small volume. This is due in part to the low volumes in Latvia of unique devices using specific apps. We also considered applying geospatial constraints restricting the search to a given neighbourhood of Riga, to better target the data points we wished to extract. Engaging with broker services revealed scattered user groups and a low availability of data in Latvia.

We contacted three different data brokers identified as serving the Latvian market and inquired about their data. From this, we struggled to find data sets with sufficient points to use for tracking, identifying individuals, or mapping behaviour; the few available sets were too expensive for the purpose of our experiment. The vendors were also reluctant to sell raw data directly to the end user and preferred to sell the data packaged or processed in some way. This does not mean that the available data is useless or cannot be exploited. It indicates only that a malicious actor will find fewer data available about a small country like Latvia, and that the available data need significant processing or aggregation to be useful, and will likely be more expensive.

Prices for data in Latvia are relatively high with regard to availability and quality. Each vendor had a starting price. The cheapest of the three we interacted with was a British-based company that quoted a base price shy of £3,000, whereas a US-based vendor gave us a base price more than six times as high. As we have already mentioned, factors such as months of historical data needed, country/region coverage required, and required data attributes play a role in determining price but this is not transparent to buyers. In our case, the available data volumes for Latvia were too low, thus, the brokers could not accurately calculate the final price.

On a positive note, the brokers we interacted with were diligent in screening potential buyers, and made inquiries about how the data would be used, stored, and processed.

The limited availability and quality of data for the Latvian market, as well as its steep price, indicate lower levels of both supply and demand compared to larger markets such as the US. With a population of less than 2 million, fewer data are generated and fewer buyers are interested in purchasing the generated data. When comparing data availability in Latvia with, for example, Kenya (see Chapter 2) it seems that market size alone does not determine price, and that the structural features of a market, such as the extent to which local businesses are data-driven, and regional legislative frameworks also influence the availability of data.

Our conclusion is that in a country or region where access to data is limited, the data marketplace and the security risks related to data brokers are smaller than in the US, where data is abundant and relatively accessible.

MARKETING ANALYSIS

To assess the potential for exploitation of the data sets available for Latvia without access to raw data, we decided to approach a marketing research company to purchase a marketing analysis; such companies essentially function as intermediaries between the data brokers and end users. The company we enlisted is based in Eastern Europe; we defined a target audience in Riga, Latvia, and asked for demographic and behavioural profiles based on mobile phone data. The data was collected for the summer of 2020, and we paid a four-figure sum for the service.

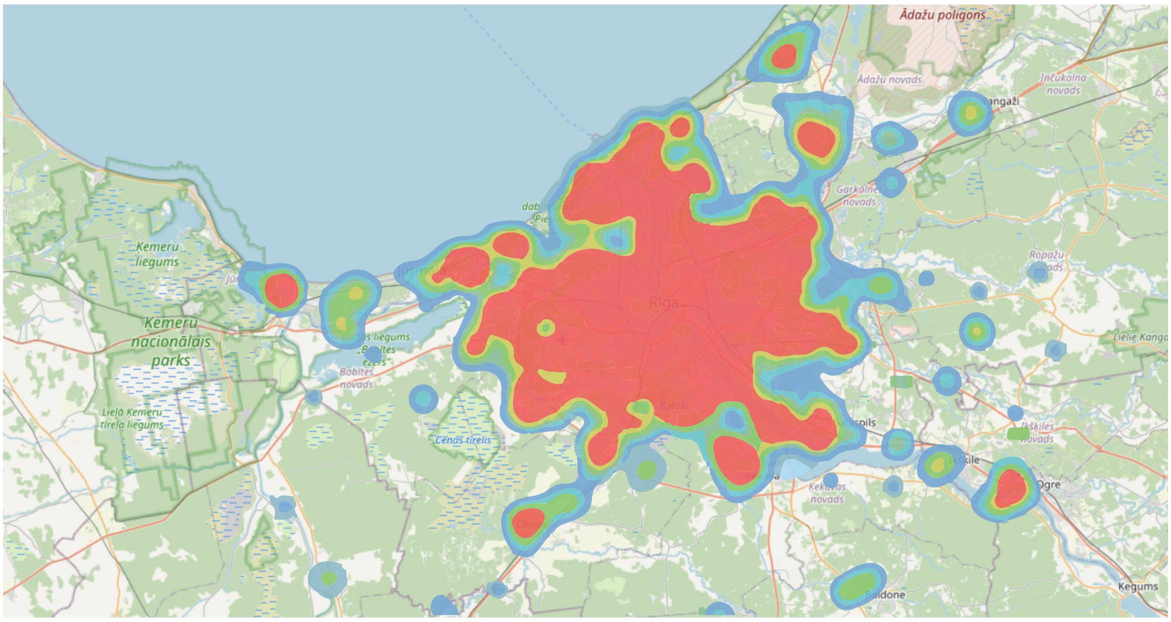
To identify the target audience, the company used unique advertisement identifiers (such as Google Advertiser Identification, or

Identifier For Advertisers) and geolocation data (GPS coordinates of users appearing within the target area at least three times during different days). In this case, the size of the identified target audience was slightly more than 100,000 individuals.

The analysis provided us with a detailed breakdown of the target audience's demographic profiles (age, gender, civil status, etc) and behavioural profiles (use of smart devices, applications, mobile banking, and coupons; patterns in shopping, travel, sports activities, and lifestyle). We could then discover where the members of our target audience shopped, which bars, coffee shops, and venues they favoured, and other points of interest. We were also given a list of the applications most used by the target audience, revealing for example which messaging applications were most popular and which websites they visited most.

The analysis also looked at geolocation data for the target audience during day-time and night-time to reveal, for example, in which areas the target audience is working or spending time during the day and where they are likely to live or at least spend time during the night.

Finally, the company provided some marketing recommendations bases the results of their analysis, for example differences between our target audience and the reference population, extrapolating a profile and suggesting appropriate channels of communication.



We used the marketing analysis product to red-team opportunities for exploitation. For example:

- Detailed demographic information in combination with geolocation data can be used to deanonymize data and identify individuals.
- Geolocation data can be used to track individuals to and from home, work, and other frequently visited locations. This could in turn be used to determine sensitive information, such as who is employed at a military facility.
- Data regarding preferred locations and venues can be exploited for surveillance and intelligence gathering, or for physical tracking.
- Device, app, and web page use could be used for malign influence campaigns or other manipulation efforts. For example, news-website preference indicates where the target audience is likely to go for new information, which could easily be exploited to mislead or deceive.

The results from our red-team analysis confirmed some of the risks identified in Chapter 4 and cleared up some questions that had arisen from our contact with the data brokers. First, although limited in comparison to other countries, the data available for Latvia are still useful; interesting data sets can be purchased at the right price and from the right vendor. Second, although raw data may be difficult to obtain, it is available to persistent buyers, and even processed data can be used for exploitation.

“Advances in the technology surrounding data harvesting and exploitation have outstripped the effect of current regulatory mechanisms.”

Leong & Yi-Ling, ‘Data Brokers: A Weak Link in National Security’, August 2020’[77]

CONCLUSION AND RECOMMENDATIONS

Our study confirms that data brokers are a weak link in national security.[78] Brokers aggregate and store huge amounts of personal data, operate with little or no transparency in an unregulated space, and are vulnerable to exploitation by malicious actors; this merits our immediate attention! For military organisations these risks are particularly important, as malicious actors can easily gather intelligence from brokers that formerly required significant resources to obtain, and exploit it to harm to critical military assets.

This report has also shown that data is not uniformly available across the globe. For a country such as Latvia, where data availability is low and the cost of quality data is high, risks are significantly smaller than for a country where data is both readily available and relatively cheap, such as the US. While our experiment demonstrated that even limited data can be exploited, it also showed that the costly and difficulty of doing so limits the potential for harm. Risk varies with data availability and capacity and resources of a malicious actor.

Vulnerabilities related to data brokers include:

- **Abundance**: Data brokers have too much data and are using it without restraint to create inferential data and to make products such as lists and personal profiles.
- **Storage**: Data is stored with insufficient security to prevent it being leaked or misused; there are no limitations on data storage so it is often retained indefinitely for potential use.
- **Use**: There is no oversight or control over how data sold by brokers is used by their customers.
- **Lack of transparency**: There is little transparency in the field of data brokarage. ‘User consent’ is glossed over; how companies collect, process, and sell their data is not at all clear.

This report concludes that efforts to mitigate the security risks related to data brokers must target all four of these aspects—through regulation, legislation, or by other means. Legislation that would limit the ability of a data broker to collect and store information that is not of immediate use or relevant for the core operation of the company could contribute to reducing risk. The EU's GDPR framework has been successful to some extent in forging a path to greater transparency within the industry,[79] but for such measures to be wholly effective users must have greater awareness of how digital data is collected and used. For example, providing the choice to accept cookies or to enable GPS features is meaningless if the user does not fully understand the consequences of that choice.

For regulatory mechanisms to be effective, brokers should be treated as part of a more complex information ecosystem within which multiple actors and institutions interact with each other and exchange data. A broader and more complete understanding of the uses and abuses of data is necessary for the creation of more accurate and meaningful measures to manage and mitigate risks in this field.

KEY TAKEAWAYS FOR NATO

To ensure mission success for Allies and Partners, the risks associated with commercially available data and data brokers should be mitigated, and the potential for exploitation by malicious actors must be limited. Based on the findings of this report, we identify five key aspects for NATO and any other military organisation to consider.

Recognize that awareness is necessary but not sufficient

A key issue is that data is generated, collected, and traded without explicit understanding and consent from the data subject. Raising awareness within military organisations about the security aspect of data and the ways in which the market for data and data brokers work is necessary as a first step towards increased digital security. Awareness in itself should not, however, be the end goal. Awareness is only useful if it leads to meaningful change in behaviour.

View data as critical infrastructure

Leong and Yi-Ling (2020) argue that today “data essentially underscores all aspects of our lives” and is a “fundamental resource for societal management and economic functioning”; it should therefore be considered critical infrastructure.[80] From a military perspective this would enable a critical infrastructure protection-approach (CIP).[81] Actions to prevent and mitigate the risks resulting from data-related vulnerabilities should include much needed changes in tactics and procedures.

Control your data

Efforts should be made to control what data is being generated by military organisations and their personnel. Such efforts could include, for example, increased training in digital security, updated SOPs that regulate use of private devices during operations, exercises, and deployments, and physical control of device use in military facilities. As we have argued elsewhere, temporarily removing phones from military personnel is not a silver bullet.[82] Continuous oversight to control, or at least limit, data generation is needed.

Red-team to understand risk

In a previous NATO StratCom COE report we used a red-teaming approach, inspired by cyber security practices, to showcase how such an approach is useful for determining the risks of open-source intelligence to military operations.[83] Red-teaming is useful for identifying risks related to commercially available data and data brokers, as we have shown here, and should be continuously applied within a military context to understand potential vulnerabilities, risks, and threats.

Leverage the potential of data

The uses for 'big data' will continue to expand, and military organisations, including NATO and its Allies, must adapt to this reality.[84] The battlefield of the future will be largely virtual in nature, and access to and insights from data will be critical in determining the outcome of conflicts.[85] Although military organisations clearly face challenges with regard to data brokers and the exploitation of commercially available data, opportunities for using data to boost capacity, improve situational awareness, and develop a tactical edge should be leveraged and not overlooked.

REFERENCES

1 Stuart A. Thompson and Charlie Warzel, "Opinion | How to Track President Trump," *The New York Times*, December 20, 2019, sec. Opinion.

2 Kale Panoho, "Council Post: The Age Of Analytics And The Importance Of Data Quality," Forbes, accessed October 7, 2020.

3 Gilles Desclaux, *Big Data & Artificial Intelligence for Military Decision Making*, 2018.

4 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

5 Sebastian Bay and Nora Biteniece, *The Current Digital Arena and Its Risks to Serving Military Personnel* (Riga: NATO Strategic Communications Centre of Excellence, 2019).

6 Bay and Biteniece; Sebastian Bay, Michael Baltra, and Henrik Twetman, Camouflage for the Digital Domain (Riga: NATO Strategic Communications Centre of Excellence, 2020).

7 Red-team analysis refers to "the practice of viewing a problem from an adversary's or competitor's perspective." Read more: "Red Teaming," Red Team Journal, accessed October 7, 2020.

8 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

9 "Global Data Broker Market – Industry Reports," 2019

10 Garrett Hazelwood, "Sell Your Data. Earn Passive Income. What Could Go Wrong?," Slate Magazine, May 28, 2019.

11 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

12 Aaron Rieke et al., "Data Brokers in an Open Society" (Open Society Foundations, 2016).

13 "Big Data: The next Frontier for Innovation, Competition, and Productivity | McKinsey," accessed October 12, 2020.

14 "Big Data: The next Frontier for Innovation, Competition, and Productivity | McKinsey," accessed October 12, 2020.

15 Gabija Fatenaite, "Web Scraping vs Web Crawling: The Differences," 2020.

16 Garrett Hazelwood, "Sell Your Data. Earn Passive Income. What Could Go Wrong?," Slate Magazine, May 28, 2019.

17 Richard Whaley, "The Big Data Battlefield—Military Embedded Systems," 2019.

18 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

19 See for example: Reuters (2018) "Facebook cuts ties to data brokers in blow to targeted ads."

20 WebFX Team and 2020, "What Are Data Brokers - And What Is Your Data Worth? [Infographic]," *WebFX Blog* (blog), March 16, 2020; Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability."

21 Acxiom, "Acxiom Launches Global Data Navigator Tool Offering Marketers Visibility into Global Audiences," Acxiom, 2018.

22 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

23 Aaron Rieke et al., "Data Brokers in an Open Society" (Open Society Foundations, 2016).

24 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

25 Forbrukerrådet, "Out of Control - How Consumers Are Exploited by the Online Advertising Industry" (Forbrukerrådet, 2020).

26 Aaron Rieke et al., "Data Brokers in an Open Society" (Open Society Foundations, 2016).

27 Kate Cox, "Secret Service Buys Location Data That Would Otherwise Need a Warrant," Ars Technica, August 17, 2020.

28 "Datarade - Find the Right Data, Effortlessly. | Datarade," accessed October 8, 2020, datarade.ai.

29 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

30 Caitlyn Renee Miller, "I Bought a Report on Everything That's Known About Me Online," *The Atlantic*, June 6, 2017.

31 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

32 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

33 "Data Brokers Are Cruising for a Bruising," *Wired*, accessed October 12, 2020.

34 "Databroker FAQ's," *Data-Broker.Co.Uk* (blog), accessed October 8, 2020.

35 Lillian Ablon, Martin C. Libicki, and Andrea A. Golay, "Characteristics of the Black Market," in *Markets for Cybercrime Tools and Stolen Data*, Hackers' Bazaar (RAND Corporation, 2014), 3–20.

36 Armor.com, "A Look inside the Dark Web - The Armor 2019 Black Market Report" (Armor, 2019).

37 Lillian Ablon, Martin C. Libicki, and Andrea A. Golay, "Characteristics of the Black Market," in *Markets for Cybercrime Tools and Stolen Data*, Hackers' Bazaar (RAND Corporation, 2014).

38 Lillian Ablon, Martin C. Libicki, and Andrea A. Golay, "Characteristics of the Black Market," in *Markets for Cybercrime Tools and Stolen Data*, Hackers' Bazaar (RAND Corporation, 2014).

39 Michael McGuire, "Into the Web of Profit - Understanding the Growth of the Cybercrime Economy" (Bromium, 2018).

40 "Data Breach Quickview - 2019 Year End Data Report" (RiskBased Security, 2019).

41 Digital shadows, "Too Much Information: The Sequel | New Research" (Digital shadows), accessed October 8, 2020.

42 Brian Stack, "Here's How Much Your Personal Information Is Selling for on the Dark Web," 2017.

43 Martin Gundersen, "Britisk dataselger varsler intern gransking etter NRK-avsløring," NRK, June 3, 2020, https://www.nrk.no/norge/mobilsporing_-britisk-dataselger-varsler-intern-gransking-etter-nrk-avsloring-1.15031158.

44 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

45 Aaron Rieke et al., "Data Brokers in an Open Society" (Open Society Foundations, 2016).

46 Daniel Miessler, "The Dark Web Has Nothing on Data Brokers," Daniel Miessler, February 1, 2020.

47 Wlosik, Michal. "What Is a Data Broker and How Does It Work? - Clearcode Blog." Clearcode | Custom AdTech and MarTech Development, February 4, 2019.

48 Stuart A. Thompson and Charlie Warzel, "Opinion | How to Track President Trump," *The New York Times*, December 20, 2019, sec. Opinion.

49 de Montjoye, Yves-Alexandre, Cesar A. Hidalgo, Michel Verleysen and Vincent D. Blondel, "Unique in the Crowd: The privacy bounds of human mobility", Nature Scientific Reports, 2013.

50 Federal Trade Commission, "Data Brokers: A Call for Transparency and Accountability," 2014.

51 "Congressional Testimony: What Information Do Data Brokers Have on Consumers? | World Privacy Forum," accessed September 21, 2020.

52 Nesita Kwan, "OfficeMax Sends Letter to 'Daughter Killed in Car Crash,'" *NBC Chicago* (blog), accessed October 8, 2020.

53 Leong, Dymples, and Teo Yi-Ling. "Data Brokers: A Weak Link in National Security." The Diplomat, 2020.

54 Grauer, "What Are 'Data Brokers,' and Why Are They Scooping Up Information About You?"

55 Paul Bischoff, "Social Media Data Broker Exposes Nearly 235 Million Profiles Scraped from Instagram, TikTok, and Youtube," *Comparitech* (blog), August 19, 2020, https://www.comparitech.com/blog/information-security/social-data-leak/.

56 Forbrukerrådet, "Out of Control - How Consumers Are Exploited by the Online Advertising Industry" (Forbrukerrådet, 2020).

57 Alexandra Ma and Ben Gillbert, "Facebook Understood How Dangerous the Trump-Linked Data Firm Cambridge Analytica Could Be Much Earlier than It Previously Said. Here's Everything That's Happened up until Now.," Business Insider, accessed October 9, 2020, https://www.businessinsider.com/cambridge-analytica-a-guide-to-the-trump-linked-data-firm-that-harvested-50-million-facebook-profiles-2018-3.

58 Pierluigi Paganini, "Genesis Store Black Marketplace Offers More than 60k+ Stolen Bot Profiles," Security Affairs, April 10, 2019,https://securityaffairs.co/wordpress/83630/deep-web/genesis-store-fingerprints.html.

59 Grauer, "What Are 'Data Brokers,' and Why Are They Scooping Up Information About You?"

60 Leong, Dymples, and Teo Yi-Ling. "Data Brokers: A Weak Link in National Security." The Diplomat, 2020.

61 Brian Krebs, "Hacked Data Broker Accounts Fueled Phony COVID Loans, Unemployment Claims — Krebs on Security," *KrebsOnSecurity* (blog), 2020.

62 NATO, "AJP-3, Allied Joint Doctrine for the Conduct of Operations (Edition C)" (NATO, 2019).

63 TAG, "Threat, Vulnerability, Risk - Commonly Mixed up Terms," *INDEPENDENT SECURITY CONSULTANTS* (blog), May 3, 2010.

64 NATO, "AJP-3, Allied Joint Doctrine for the Conduct of Operations (Edition C)."

65 Försvarsmakten, "Handbok Säkerhetstjänst Grunder" (Försvarsmakten, 2013).

66 NATO, "AJP-3, Allied Joint Doctrine for the Conduct of Operations (Edition C)."

67 Thompson, Stuart A., and Charlie Warzel. "Opinion | How to Track President Trump." The New York Times, December 20, 2019, sec. Opinion.

68 Thompson, Stuart A., and Charlie Warzel. "Opinion | How to Track President Trump." The New York Times, December 20, 2019, sec. Opinion.

69 Foeke Postma, "Military And Intelligence Personnel Can Be Tracked With The Untappd Beer App," bellingcat, May 18, 2020.

70 Alex Hern, "Fitness Tracking App Strava Gives Away Location of Secret US Army Bases," *The Guardian*, January 28, 2018, sec. Technology.

71 Mike Elgan, "5 Shocking New Threats to Your Personal Data," Computerworld, February 4, 2017.

72 Gundersen, Martin, "Norske offiserer og soldater avslørt av mobilen." NRK, May 18, 2020.

73 Gundersen, Martin. "Britisk dataselger varsler intern gransking etter NRK-avsløring." NRK, June 3, 2020.

74 Thompson, Stuart A., and Charlie Warzel. "Opinion | How to Track President Trump." The New York Times, December 20, 2019, sec. Opinion.

75 Red Team Journal. "Red Teaming." Accessed October 7, 2020.

76 "Datarade - Find the Right Data, Effortlessly. | Datarade." Accessed October 8, 2020. datarade.ai.

77 Leong, Dymples, and Teo Yi-Ling. "Data Brokers: A Weak Link in National Security." The Diplomat, 2020.

78 Leong, Dymples, and Teo Yi-Ling. "Data Brokers: A Weak Link in National Security." The Diplomat, 2020.

79 Leong, Dymples, and Teo Yi-Ling. "Data Brokers: A Weak Link in National Security." The Diplomat, 2020.

80 Leong, Dymples, and Teo Yi-Ling. “Data Brokers: A Weak Link in National Security.” The Diplomat, 2020.

81 NATO, “Critical Infrastructure Protection,” NATO, 2020.

82 Bay, Sebastian, Michael Baltra, and Henrik Twetman. “Camouflage for the Digital Domain.” NATO StratCom Centre of Excellence, 2020.

83 Bay, Sebastian, and Nora Biteniece. “The Current Digital Arena and Its Risks to Serving Military Personnel | StratCom.” NATO StratCom Centre of Excellence, 2019.

84 Jean Brunet and Nicolas Claudon, “Chapter 7 - Military and Big Data Revolution,” in *Application of Big Data for National Security*, ed. Babak Akhgar et al. (Butterworth-Heinemann, 2015), 81–107, https://doi.org/10.1016/B978-0-12-801967-2.00007-0.

85 Whaley, Richard. “The Big Data Battlefield - Military Embedded Systems,” 2019.

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