

Review report

Automated OSINT: tools and sources for open source investigation

CTIVD no. 74

Adopted on 22 December 2021



Review Committee
on the Intelligence and
Security Services

Disclaimer: This is not an official translation. No rights may be derived from this translation and under all circumstances the Dutch text of this report prevails.

Inhoudsopgave

Summary	3
1. Introduction	7
2. Automated OSINT: Examples of tools and sources	11
2.1 Types of OSINT	11
2.2 Development in OSINT	15
2.3 Specialized tools and commercially available sources	15
2.4 Selection of tools by the services	16
2.5 The use of automated OSINT tools at the AIVD	17
2.6 The use of automated OSINT tools at the MIVD	18
2.7 Differences between automated OSINT at the AIVD and the MIVD	18
3. Automated OSINT: Review of the legal framework	19
3.1 General provisions regarding data processing	19
3.2 Proper and careful data processing	19
3.3 Reliability and accuracy of data	21
3.4 Automated data analysis	21
3.5 Compliance with the duty of care for source and identity of staff	22
3.6 Compliance with the duty of care regarding data processing	22
4. Conclusions and recommendations	24

Summary

The Intelligence and Security Services Act 2017 (ISS Act 2017) permits the General Intelligence and Security Service (AIVD) and the Military Intelligence and Security Service (MIVD) to collect and process publicly accessible data, including personal data. These activities are also referred to as open source investigation or OSINT, which stands for 'open source intelligence'. The Dutch legislator does not consider OSINT to be an intrusive intelligence method. It is regarded as a general investigatory power of the services, whereby a distinction is made between non-systematic and systematic deployment. Authorization is required for the systematic collection of personal data from information sources accessible to the general public.

When open source investigation is automated using specialized software or web applications, it is referred to as 'automated OSINT'. This investigation makes a distinction between the tools that are used and the sources (data sets) that may be accessed through these tools. The tools are in fact software equipped with search and network analysis functions which can query a wide variety of sources. These tools may come from commercial providers or be developed by the services themselves.

Tools for automated OSINT offer two major advantages over standard open source investigation using a web browser. The first of these is ease of use: a single search using an automated OSINT tool can query hundreds of sources simultaneously. The tool can then provide a visual representation of the results. The second major advantage to the services of using those tools is that they give access to sources based on user-friendly services provided by the tool's vendor on a commercial basis. One example of this is leaked data from users of social media services. Vendors can aggregate these data sets as a single searchable source (a 'composite data set'), which may contain billions of data points.

One example of commercial data that can be accessed through these tools is location data generated by ads shown to application users. Providers of commercial tools for OSINT can purchase advertisement data from data brokers and use their tool to make it available to clients, including intelligence and security services.

The volume, nature and types of personal data in these automated OSINT tools may lead to a more serious violation of fundamental rights, in particular the right to privacy, than consulting data from publicly accessible online information sources, such as publicly accessible social media data or data retrieved using a generic search engine.

From the explanatory memorandum of the ISS Act 2017, it can be concluded that the practices facilitated by automated OSINT were not taken into account by the legislator at the time and will continue to develop in the near future. OSINT undeniably goes well beyond investigative techniques such as checking telephone directories or using a search engine to access online data. The present investigation reflects the following reality: automated OSINT provides simultaneous searchable access to hundreds of sources of various origins, including location data or data from leaked data sets. The current practice of automated OSINT involves a more serious violation of privacy than was anticipated at the time.

This finding leads to **recommendation 1**:

Given the nature, diversity, and volume of the data at issue, the Review Committee on the Intelligence and Security Services (CTIVD) recommends that the legislator creates a more foreseeable legal basis with sufficient safeguards governing the use of automated OSINT, both the tools themselves and the sources that can be accessed using these tools.

In the present review, the CTIVD's lawfulness assessment focuses primarily on the OSINT tools and the data sets (i.e. the sources) that can be accessed using these tools. How these tools and sources are used in actual cases does not fall within the remit of the review. The CTIVD considers it essential that the services know how the tools work and which sources can be consulted prior to actual deployment. Only with this knowledge can a thorough assessment be conducted to determine how the processing of this data with these tools relates to the general data processing provisions of the ISS Act 2017. Among other things, these provisions require that the processing of data by the services should be proportionate. This means that there must be an appropriate balance between the interests at stake in processing the data for the relevant intelligence investigation and the severity of the breach of the fundamental rights of the data subject.

By conducting this review, the CTIVD aims to answer the following research question:

Do the AIVD and the MIVD have a sufficient understanding of the workings of the automated OSINT tools and the origin and the nature of the underlying sources with a view to complying with the data processing provisions?

The answer to the investigative question is that the AIVD's and the MIVD's understanding of the workings (the functionalities) of the automated OSINT tools and the origin and nature of the sources that can be consulted using these tools is insufficient to ensure compliance with the data processing provisions of the ISS Act 2017. The CTIVD notes that several improvements are needed before automated OSINT can be brought into compliance with the law. The services should identify the workings and (as thoroughly as possible) the underlying sources of the tools and take mitigating measures in this regard to prevent unlawful conduct in the future.

This finding leads to **recommendation 2**:

When selecting and acquiring tools for automated OSINT (and thereby selecting the underlying sources), the AIVD and the MIVD should also aim to ensure lawful data processing. Preferably, the services should work together to develop a joint policy framework with accompanying work instructions.

In the interests of legal certainty, lawfulness and operational effectiveness on the part of the services (based on the continuity of lawful data processing using OSINT), the CTIVD will enter into a dialogue with the services in order to arrive at a workable temporary assessment framework which the services will then translate into policy, procedures, and work instructions. This temporary assessment framework should, among other things, address the establishment of a prior assessment in light of the data-processing provisions, the criterion of a systematic approach to open source investigation, and the handling of sources whereby the origin and accuracy of the data cannot be clearly established.

The use of OSINT is not exclusive to the domain of the intelligence and security services, but also applies elsewhere in the national security domain (for example, at the National Coordinator for Security and Counterterrorism) and beyond (including other government bodies). This review report notes that OSINT has continued to evolve over the years, allowing for the use of tools that simultaneously consult hundreds of sources. The results of this process can be displayed rapidly, clearly and in context. These underlying sources may include location data or leaked data. The processing of such data constitutes a violation of the fundamental rights of data subjects that goes further than OSINT using standard search engines or social media services.

The CTIVD therefore asks the Minister of the Interior and Kingdom Relations and the Minister of Defence to bring this report to the attention of other government bodies and, when forwarding the report, to ask Parliament to bring it to the attention of the House of Representatives' Standing Committee on Digital Affairs.

1. Introduction

This review report concerns automated open source intelligence (hereinafter: automated OSINT). In this report, OSINT is also referred to as 'open source investigation' and with the term used in the ISS Act 2017 'the collection of data from publicly accessible information sources'. Automated OSINT is conducted with the use of tools such as software or web applications. The tools may come from commercial providers or be developed by the General Intelligence and Security Service (AIVD) and the Military Intelligence and Security Service (MIVD) themselves.

The tools for automated OSINT on which this investigation focuses, contain search and analytical functions with access to a wide range of sources such as articles on news websites and publicly accessible data on social media services. These tools can also be used to consult personal data that has been collected and pre-processed by a commercial provider. These 'composite data sets' may contain leaked personal data from users of social media services. Furthermore, these tools open the door to consulting commercially available data, such as a data set with location data.

AIVD and MIVD employees with access to the tools can enter a query into the tool, for example the question whether a profile on a social media service can be linked to a target (a person who is the subject of the services' focus) and thereby collect personal data on that target. This is referred to as target-oriented investigation. The tools can also be used to conduct a phenomenological investigation, that looks in particular at social developments and identifies trends. Examples include collecting and analysing information from news articles to obtain an overview of a certain region or of a certain theme, which could mean that personal data is processed. Given the fact that target-oriented investigations using tools for automated OSINT generally constitute a greater infringement of privacy than phenomenological investigations, this current review report focuses in particular on target-oriented investigation.

Legal regulation

Section 25 of the Intelligence and Security Services Act 2017 (ISS Act 2017) provides the AIVD and the MIVD with the investigatory power to collect data from publicly accessible information sources. Based on Section 38 of the ISS Act 2017, the services are also authorized to *systematically* collect data on people from publicly accessible information sources, with or without the use of a technical tool. Systematic collection requires an application to use the special investigatory power and authorization must be granted.¹

¹ See also the assessment framework to this report (Appendix I).

The distinction between open-source investigation and systematic open-source investigation was made by the legislator in the ISS Act 2017 following the Privacy Impact Assessment ISS Act 20XX. That report showed clearly that systematic open-source investigations into individuals can lead to a more severe breach of privacy.² An investigation is deemed to be systematic if it can reasonably be foreseen in advance that a more or less complete view of certain aspects of an individual's private life will be obtained.

The data processed by the services is subject to the general provisions regarding data processing in the ISS Act 2017. Those provisions prescribe that the data is processed for the services' tasks, that it is proportional and that the head of the service is responsible for taking measures to improve the accuracy and completeness of the data to be processed and to improve the quality of the data processing.³ A further explanation to the legal framework can be found in the assessment framework (Appendix I).

Public debate

The social impact of open source investigations within the national security domain again became clear in 2021 during the debate on the use of fake profiles on social media services, raised by the National Coordinator for Security and Counterterrorism (NCTV).⁴ There is also increasing focus on OSINT abroad. In the United States, for example, the use of the tool Locate X by US government services led to questions from US senators and a pending investigation by the Office of Inspector General of the Department of Homeland Security.⁵ Furthermore there is currently an ongoing investigation into the FBI Collection of Open Source Data by the independent oversight body Privacy and Civil Liberties Oversight Board (PCLOB).⁶ In the United States the debate focuses mainly on the question if a judicial authorization – or warrant – is necessary in requests for location data using those tools.

The implications of open source investigations are also attracting attention beyond the national security domain.⁷ A bill is currently being prepared in the Netherlands which lays down rules for an identical investigatory power within the crime investigation domain governing the systematic copying of personal data from publicly accessible information sources.⁸

Scope of the investigation and investigative question

The CTIVD's lawfulness assessment focuses primarily on the OSINT tools and the data sets (i.e. the sources) that can be accessed using these tools. How these tools and sources are used in actual cases and the question if that use is systematic does not fall within the remit of the review. This investigation looked at whether, in the context of lawful data processing, advance thought was given to compliance with the general provisions for data processing in the ISS Act 2017. The investigation

² B.J. Koops et.al., 'Privacy Impact Assessment Wet op de Inlichtingen- en veiligheidsdiensten 20XX' ('Privacy Impact Assessment Intelligence and Security Services Act 20XX'), TNO/PILab/Tilburg University, 2016. *Parliamentary Documents II 2016/17*, 34588, no. 3, p. 63 and *Parliamentary Documents II 2016/17*, 34588, no. 18, p. 53.

³ Sections 18 and 24 of the ISS Act 2017, respectively.

⁴ See A. Kouwenhoven, E. Rosenberg & R. van der Poel, 'NCTV volgt heimelijk burgers op sociale media' ('NCTV covertly tracks citizens on social media'), *NRC*, 9 April 2021 and A. Kouwenhoven, E. Rosenberg & R. van der Poel, 'Linkse activist werd jaren online gevolgd door de NCTV' ('Left-wing activist was tracked by the NCTV for years online'), *NRC*, 23 July 2021.

⁵ 'DHS Use of Cell-Phone Surveillance Devices', *Office of Inspector General, U.S. Department of Homeland Security*, accessible on www.oig.dhs.gov/node/6227. See also, among others, B. Tau, 'U.S. Government Contractor Embedded Software in Apps to Track Phones', *The Wall Street Journal*, 7 August 2020, J. Cox, 'Secret Service Bought Phone Location Data from Apps, Contract Confirms', *Vice.com*, 17 August 2020, 'How the U.S. Military Buys Location Data from Ordinary Apps', *Vice.com*, 16 November 2020 and C. Savage, 'Intelligence Analysts Use U.S. Smartphone Location Data Without Warrants, Memo Says', *The New York Times*, 22 January 2021.

⁶ 'FBI Collection of Open-Source Data', U.S. *Privacy and Civil Liberties Oversight Board*. Accessible www.pclob.gov/projects.

⁷ See among others, H. von Piekartz, 'Gemeenten kijken op grote schaal en in het geheim mee met burgers op sociale media' ('Municipalities are monitoring citizens covertly and widely on social media'), *De Volkskrant*, 18 May 2021 and E. Rosenberg & K. Berkhout, 'Leger verzamelde data in Nederland' ('The army collected data in the Netherlands'), *NRC*, 15 November 2020.

⁸ Bill to adopt the new Code of Criminal Procedure, July 2020 (official version)

focuses predominantly on the tools – how are these selected, how do they work and how are the general requirements for data processing observed? Secondly the investigation looks at the sources consulted using those tools: what are those sources and how do they relate to the general provisions on data processing?

Thereby the CTIVD aims to answer the following research question:

Do the AIVD and the MIVD have a sufficient understanding of the workings of the automated OSINT tools and the origin and the nature of the underlying sources with a view to complying with the data processing provisions?

The investigation period selected by the CTIVD covers the period from 1 June 2020 to 31 March 2021. The services indicated, on being asked, that this period is representative.

The decision to limit this investigation to the preliminary phase was prompted by the fact that it appeared during the exploratory investigation that the services had given too little thought before and during the acquisition process of a tool to comply with the general requirements for data processing when using automated OSINT. Unlawful conduct is therefore highly likely when using tools for automated OSINT to collect and further process data. Furthermore, open source investigations based on the general investigatory power under Section 25 of the ISS Act 2017 by the services did not always appear to be logged or recorded if they failed to yield results. Oversight of compliance with Section 25 of the ISS Act 2017 in relation to Section 38 of the ISS Act 2017 (the systematic collection of data on a person from publicly accessible information sources) is therefore not possible.

The CTIVD feels it is essential for the services to set up the process of automated OSINT in line with the ISS Act 2017. To do so, the services will need to fully understand how the tools function, such as the functionalities that the tools offer and the underlying sources. That knowledge will help them learn how the tools and sources relate to the legal requirements for data processing and the legal investigatory power to conduct systematic or other open source investigations.

The underlying data that can be consulted through the tools for automated OSINT can take the form of a bulk data set. That aspect as such is not a subject of this investigation, but was addressed in CTIVD report 55 (on the acquisition of bulk data sets offered on the internet by third parties), report 70 (collecting bulk data sets using the hacking power and their further processing) and report 71 (the collection and further processing of passenger data from airlines).

Methodology

The tools for automated OSINT and the underlying sources are assessed on lawfulness using the assessment framework (Appendix I). An 'unlawful' assessment always means that the conduct conflicts with legislation and regulations. Legislation and regulations in this case refers to the ISS Act 2017, case law and the recommendations adopted by ministers based on previous CTIVD review reports. Any deficiencies in the procedures, policies or processes of a service are referred to as negligence in the report.

During the investigation, interviews were conducted with staff of various teams at both services. Internal documents at the AIVD and the MIVD were studied, such as policy and work instructions relating to open source investigation and work documents relating to the tools used for automated OSINT. Furthermore, literature and commercial (OSINT) products were studied to obtain the necessary understanding of the tools. The services confirmed that they submitted all information relevant to answering the investigative question.

Classified appendix

This report has a classified appendix. This appendix does not contain any reports of unlawful conduct that have not been described in the public review report. However, the classified appendix contains more detailed information that reveals the services' procedure relating to automated OSINT and for that reason had been marked 'classified'.

Structure of the report

The report has the following structure. Section 2 provides an outline of the different ways in which OSINT and automated OSINT is conducted in general and within the AIVD and the MIVD. Section 3 lists the findings on compliance with legal requirements for data processing in automated OSINT practice at the AIVD and the MIVD. The report concludes with section 4 which describes the conclusions and recommendations. The report has three appendices – a classified appendix, an assessment framework (Appendix I) and the definitions (Appendix II).

2. Automated OSINT: Examples of tools and sources

OSINT can be conducted in various ways and the field of OSINT has developed over the past years into a high-tech intelligence method. However, for the remainder of the report it is important to bear in mind that there are now, in 2021, significantly more options for collecting and processing data using specialist tools than in 2014, when the CTIVD conducted an investigation into the collection by the AIVD of data from social media services.⁹ These specialist tools can aid the services in consulting various social media services – such as Facebook and Twitter – simultaneously, but large amounts of commercially available data can also be used. The explanatory memorandum to the ISS Act 2017 does not refer to this development in the field of OSINT, so it is unclear to the general public to what extent an infringement of fundamental rights can occur as a result of those developments.

This section explains how OSINT and automated OSINT work in general and how the tools for automated OSINT are applied at the AIVD and the MIVD. Firstly, we look at the various forms of OSINT (including the use of tools) and subsequently at the sources that can be consulted with those tools (sections 2.1 and 2.2). Section 2.3 describes the development of specialized tools and commercially available sources. The AIVD and MIVD's practice in terms of automated OSINT is discussed in sections 2.4 to 2.7.

2.1 Types of OSINT

This section provides an overview of the development in the field of OSINT and automated OSINT. OSINT can be divided into three categories, namely consultation (1) through a web browser or app, (2) through an Application Programming Interface (API) and (3) by using a specialized application (tool).

⁹ Review report no. 39 (2014) on investigation of social media by the AIVD.

OSINT through a web browser or app

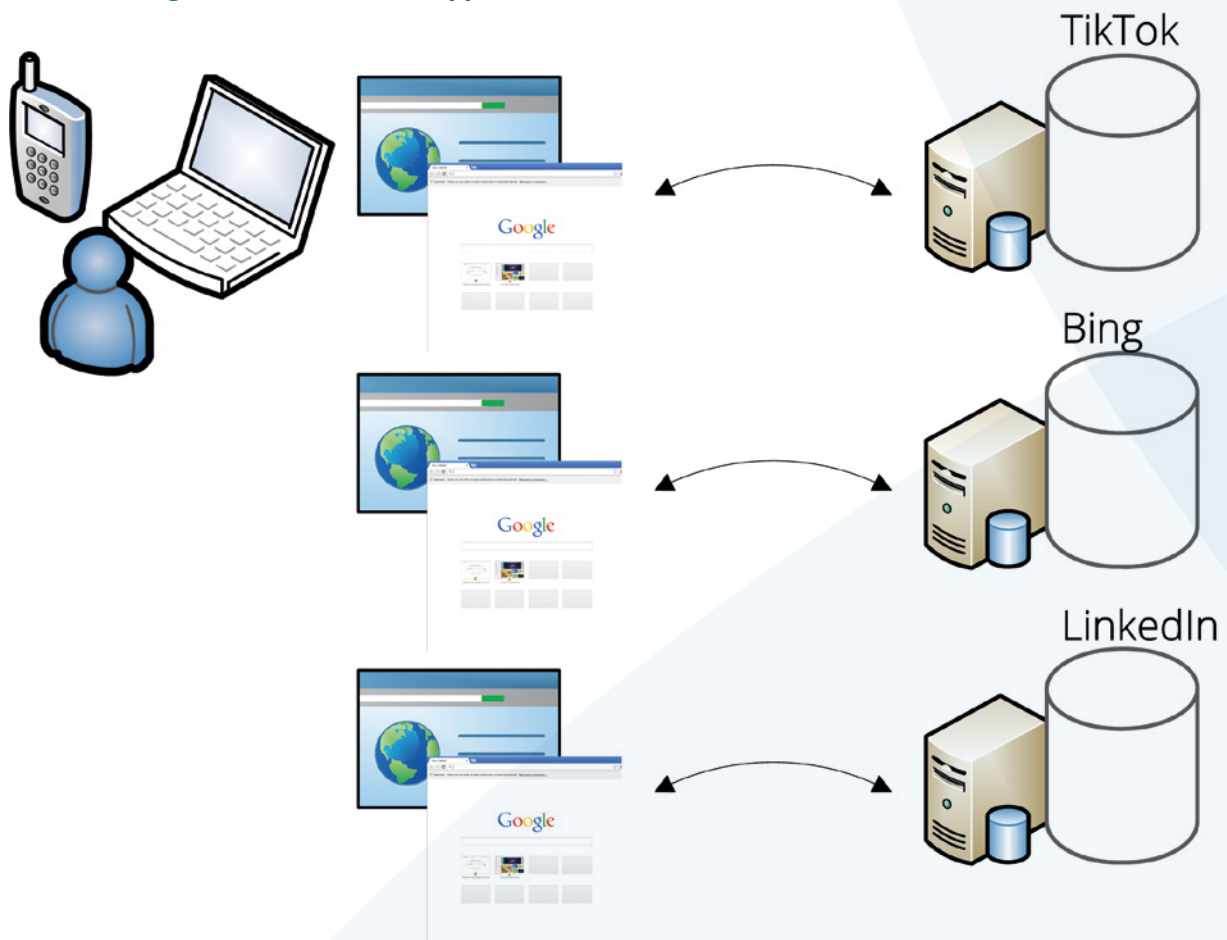


Figure 2.1: OSINT through the web browser with some sample sources.

Everyone is familiar with search engines such as Google or with browsing a profile on a social media service such as Facebook or LinkedIn using a web browser – for example Google Chrome, Firefox or Microsoft Edge – on a laptop or an app on a mobile phone.

In general, search engines index web pages that are available on the internet. These are collected, indexed and stored in a database by crawlers. That database can be searched using a search engine such as Google. Other common search engines are Bing, Baidu and Yandex. Apart from the obvious search engines, there is such as thing as The Wayback Machine, a digital world-wide-web archive.¹⁰ Apps generally provide access to a specific platform, have an independent function or complement search engines.

¹⁰ See among others, H. Gibson, 'Acquisition and Preparation of Data for OSINT Investigations', in: B. Akhgar, S. Bayerl & F. Sampson, *Open Source Intelligence Investigation*, Springer: 2016.

Using an API

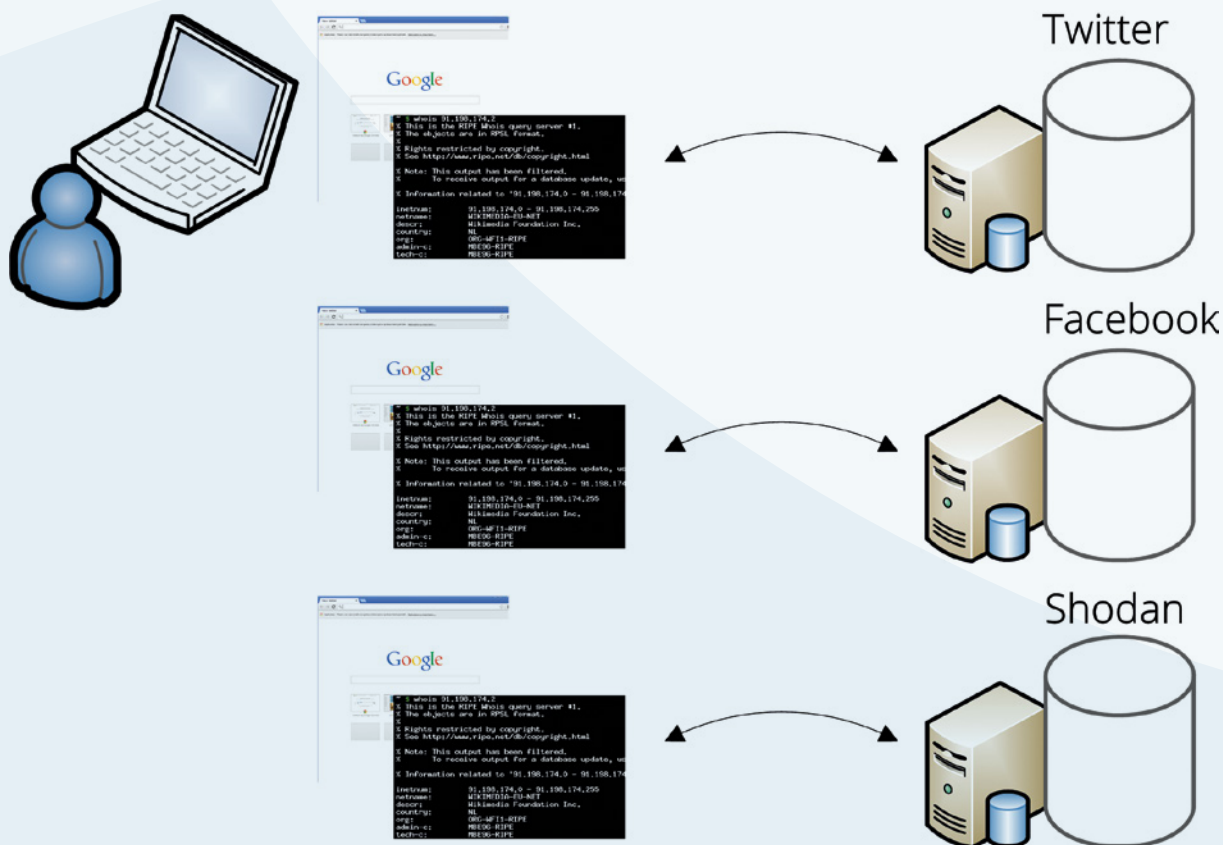


Figure 2.2: OSINT through API with some sample sources.

An Application Programming Interface (API) is a collective term for a computer program that communicates with another program (or: backend) to enable an exchange of data. One example of an API is the functionality offered by many websites to log onto their platform using a social media account such as Facebook, LinkedIn or Google.

The API provides the website with identification data that authorizes the user to access the underlying social media platform. Sometimes an account must be created in order to access the API. The advantage of consulting data via an API as opposed to visiting a webpage using a browser, is that it is easier to retrieve structured data without unnecessary overhead information such as the layout of the web page. Another important distinction is that an API search is easier to automate. That makes open source investigation more efficient compared to a web browser.

Tools for automated OSINT

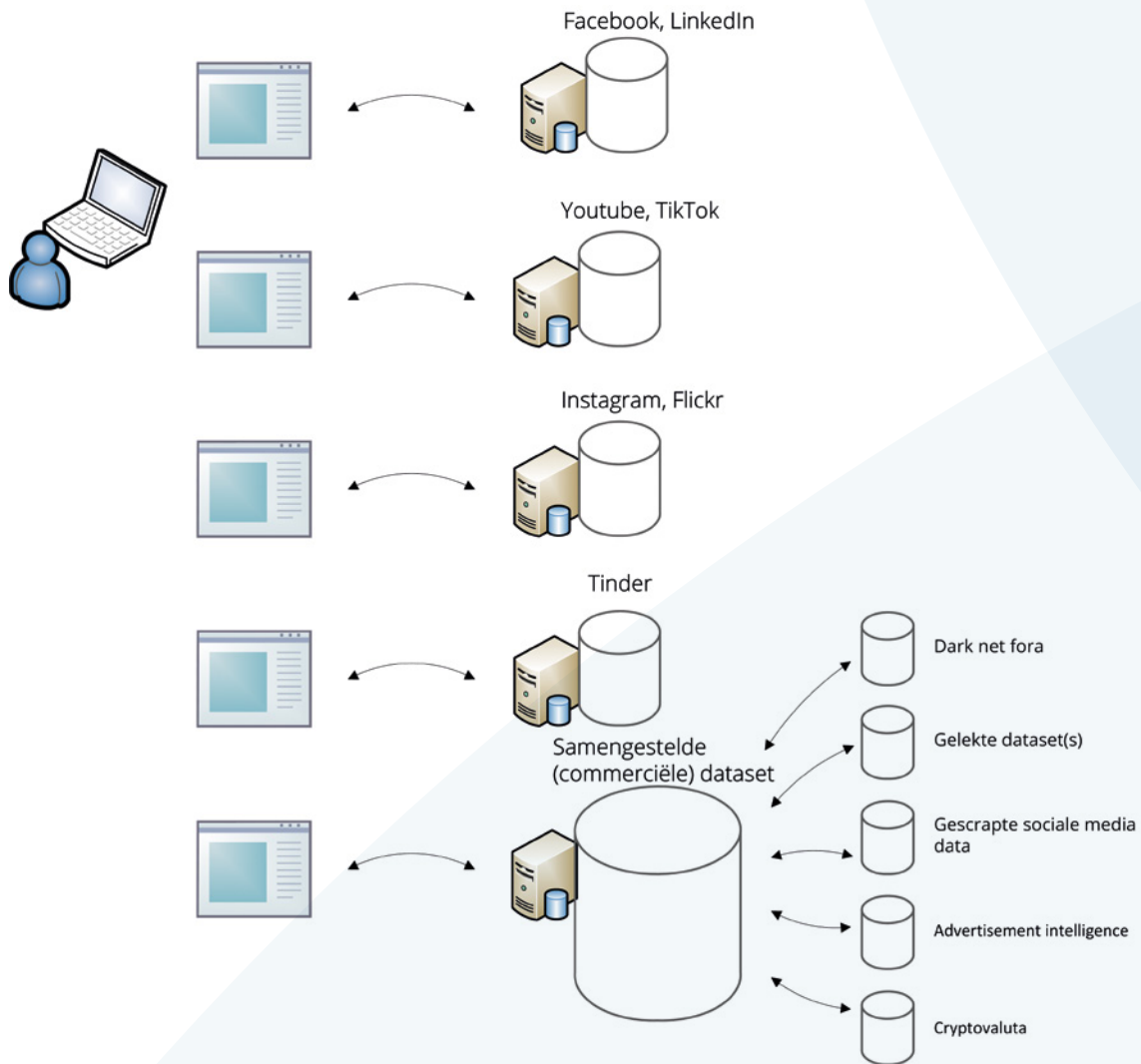


Figure 2.3: OSINT and automated OSINT through tools with some sample sources.

Tools for automated OSINT combine the use of OSINT via the web browser with data processing through the API of websites and data processing in a 'composite' commercial data sets. That makes it possible to consult all the above sources simultaneously using a single search criterion, such as the name of a person or a phone number. The results are then made available to the user of the tool, in context or otherwise. This is an efficient way of consulting a large amount of data sources to obtain an overview of the available information about an individual.¹¹ Consulting the same data from different sources can contribute to the accuracy and reliability of that information. The data of interest to the investigation is then copied and processed further for the execution of tasks.

The methods to analyse data using these tools are divided into three categories in the literature, i.e. lexical analysis, network analysis and geo-spatial analysis or a combination of these three variants.¹² Lexical analysis refers to the collection of large amounts of texts. The tools also enable network analyses and visualizations, helping to throw light on relationships between individuals or entities, for example. Finally, geo-spatial analysis is applied to link certain data to a specific location (also known as geo-tagging). The available tools then make it possible to visualize and analyse that data and reveal possible displacements of individuals and objects.

¹¹ See among others, H. Gibson, 'Acquisition and Preparation of Data for OSINT Investigations', in: B. Akhgar, S. Bayerl & F. Sampson, *Open Source Intelligence Investigation*, Springer: 2016.

¹² H. Williams and I. Blum, *Defining Second Generation Open Source Intelligence (OSINT) for the Defense Enterprise*, RAND: 2018, p. 23.

2.2 Development in OSINT

In OSINT the data referred to is traditionally derived from phone books, newspapers, journals, news and other websites, analyses by think tanks or scientific articles. These sources generally have static content published by organizations, press agencies, governments and individuals.

However, data from the deep web or 'dark web' such as online forums or darknet markets, can be publicly accessible information. The deep web is a part of the internet which is not indexed by the usual search engines for the regular internet, but which is in fact accessible if the website is visited. The dark web consists of a section of the internet where the IP addresses of networked computers are hidden using special software, such as The Onion Router (Tor).

In this day and age, user-generated content (information supplied by users of a certain medium, in particular social media) plays an important part in finding personal data.¹³ That can also include sharing messages (as on Facebook and LinkedIn), sharing videos (as on YouTube and TikTok) and sharing photos (as on Instagram and Flickr). In addition, there is a brisk trade in the data of visits by individuals to websites and their activities on social media services. Information from adverts, but also from cookies, are traded by a range of parties and may eventually also end up with other parties, such as the intelligence and security services through a company offering tools for automated OSINT.

2.3 Specialized tools and commercially available sources

Sources are increasingly being made available commercially or becoming accessible, whether or not through specialized tools. There are services on the market who supply different types of products and who facilitate the collection and processing of data from open sources. That includes packages such as SpiderFoot, Recon-ng, Maltego or Spysse.¹⁴ It is this type of tools that could be the subject of this report.

Some tools contain modules or plug-ins for various OSINT applications. Those could be free or paid modules or plug-ins. With the help of a plug-in or module, a user can automatically search a source such as Facebook or LinkedIn with a couple of mouse clicks. These plug-ins or modules can also consist of aggregated data sets. The revenue model of these companies is based on offering paid modules or plug-ins.

Commercial parties also offer specialized data sets that they have aggregated themselves. These data sets sometimes contain data 'scraped' from the internet, meaning that the data is automatically collected from publicly accessible sources based on prefixed parameters.¹⁵ The providers of these data sets may be reluctant to provide information about the origin of those data sets and how the data sets are structured. Information about users generated by the providers of apps or websites (e.g. via cookies or adverts) are often sold to or exchanged with other parties. That data can then end up in a commercial data set which may be searched using a tool for automated OSINT.

¹³ See also review report no. 39 (2014) on the use of data from social media and internet forums by the AIVD.

¹⁴ See www.tools.kali.org/information-gathering/recon-ng; www.maltego.com; and www.spysse.com. SpiderFoot is described as follows: "a reconnaissance tool that automatically queries over 100 public data sources (OSINT) to gather intelligence on IP addresses, domain names, e-mail addresses, names and more" (<http://spiderfoot.net/documentation/>).

¹⁵ See among others, H. Gibson, 'Acquisition and Preparation of Data for OSINT Investigations', in: B. Akhgar, S. Bayerl & F. Sampson, *Open Source Intelligence Investigation*, Springer: 2016.

The data sets offered by these parties may consist of the following:

- Scraped historical data: that involves data from multiple social media sources, but many commercial parties also scrape data from internet forums or market places on the dark web;
- Crypto currency: information about crypto currency transactions (tokens, addresses and users);
- Leaked data sets: sets consisting of data obtained through a hack. Leaked data sets may have been made available through open sources or through a commercial party only;
- Advert data: data sets using location data collected by purchasing advertising space in an app or on a website.¹⁶

Providers of tools for automated OSINT collect that data, process it further and offer it as data source to customers, such as the intelligence and security services. Data from different data sets containing personal data may be processed by the provider and aggregated to a single data set. A 'composite' data set may contain billions of data points.

The data sets offered commercially and which are accessible by specialized tools include advertisement-based intelligence (ADINT). This is data from mobile devices, such as location data and data about the mobile device. That data is generated by advertisements in apps and the data is then traded on data and other markets. The providers of that data are referred to as 'data brokers'.

In so far as commercially available data cannot be viewed as a publicly accessible information source, the services are authorized under Section 25(1)(b) of the ISS Act 2017 to process commercially available data to execute their tasks.¹⁷ The CTIVD notes that the scope of that investigatory power is not clear from the section of the law because it refers to 'information sources for which the service has been granted a right of access the data stored within.' A 'right of access' can be apparent from a basis in other legislation, such as the Police Data Act, on which grounds the AIVD and the MIVD may consult the data held there. The explanatory memorandum to the ISS Act 2017 only makes one reference to the collection of commercially available data by the services based on the section of the law and gives no example, other than information from the Chamber of Commerce.¹⁸ More clarity about the scope of this legal basis for processing commercially available data and appropriate legal safeguards are therefore advisable.

2.4 Selection of tools by the services

When an organization decides to deploy automated OSINT tools, it may opt to develop those tools itself (or have them developed). That has the advantage of creating a tailor-made tool of which all functionalities and underlying sources are known. The downside is that this requires specialized knowledge, development time and money. A ready-made market product will not have those drawbacks but will have other disadvantages, for example that the exact working of the tool is unclear and that only limited influence can be exerted on functionalities such as search options and logging. In certain cases, data sets cannot be purchased without the tool because that is part of the revenue model of the product.

Decisions on purchasing tools for automated OSINT therefore call for a careful weighting process. If the AIVD or the MIVD are interested in a certain product in the field of automated OSINT, the product is not first assessed on previously established criteria within the service.

¹⁶ P. Vines, F. Roesner and T. Kohno, 'Exploring ADINT: Using Ad Targeting for Surveillance on a Budget - or - How Alice Can Buy Ads to Track Bob', *The 16th ACM Workshop on Privacy in the Electronic Society (WPES 2017)*. See also H. Gibson, 'Acquisition and Preparation of Data for OSINT Investigations', in: B. Akhgar, S. Bayerl & F. Sampson, *Open Source Intelligence Investigation*, Springer: 2016.

¹⁷ Section 25(1)(b) of the ISS Act 2017. See also *Parliamentary Documents II 2016/17*, 34588, no. 3, p. 38. See also Section 1.2 of the Assessment Framework (Appendix I).

¹⁸ *Parliamentary Documents II 2016/17*, 34588, no. 3, p. 38.

The investigation shows that the services devoted a great deal of attention to the possible added value of the tools compared with other options for open source investigation. They looked carefully at the purpose of the remedy and at which data is collected. The primary focus was on the operational added value of the specific product and the operational security, also of staff, (see section 3.5) and less on the breach that the deployment of the tool could mean for the fundamental rights of the data subjects.

In the context of the duty of care for lawful data processing (Section 24 of the ISS Act 2017), attention should also be devoted to aspects of compliance such as dash boarding and logging. Due care of data processing (see section 3.2) must be included in an assessment. Obtaining a picture of those aspects will contribute to a sound level of knowledge of the tools' functioning and the nature of the underlying data sets. That knowledge is necessary when weighing the proportionality of the data processing using the tools. In addition, that knowledge is necessary to determine which measures should be taken to mitigate risks in light of the duty of care for data processing and to achieve internal control with effective oversight.

2.5 The use of automated OSINT tools at the AIVD

During the investigation period from 1 June 2020 to 31 March 2021, the AIVD made very little use of tools for automated OSINT. The use of those tools was reserved for a number of designated officials.

During the investigation period, only one AIVD team that focuses on identifying new targets threatening national security, gained some experience with the use of an externally purchased tool for automated OSINT. The aim was to see whether the tool had any added value for the operational practice. An internal evaluation by the AIVD revealed that the automated OSINT tool did in fact have added value, because the automatic search of many sources simultaneously made their investigation more efficient. Staff are able to collect more data on an individual in less time. In the investigation period (and up to the adoption of this report) the tool was still in the test phase. However, the tool was deployed twice in the operation process during the investigation period. To that end, the investigatory power based on Section 38 of the ISS Act 2017 for the systematic collection of personal data from publicly accessible information sources was used only a single time.

In addition, data analysts from the service used a second, externally purchased tool for automated OSINT to process data about targets. It is one of the tools that data analysts have at their disposal to carry out their tasks. During the investigation period, the tool was deployed several times based on the general investigatory power under Section 25 of the ISS Act 2017.

The AIVD also has a department of specialized staff who conduct more in-depth open source investigations at the request of a team. That often means target-oriented investigation, where data is collected about an individual or an organization who is the focus of the service. That team did not use any tools for automated OSINT at the time of this investigation.

Finally, staff with certain positions in almost all AIVD teams conduct open source investigations to a certain depth without the use of tools for automated OSINT.

The limited use of tools for automated OSINT can be explained by the fact that during the investigation period, the AIVD mainly tested whether tools for automated OSINT had any added value for the organization. Thus, the tools are not, or not yet, part of a standing process for collecting data through open source investigation by staff of teams or by the specialist staff members of the OSINT department.

2.6 The use of automated OSINT tools at the MIVD

During the investigation period from 1 July 2020 to 31 March 2021, the MIVD made frequent use of tools for automated OSINT.

The tools for automated OSINT are deployed by a specialist department of the MIVD and its use is a standard component of the intelligence process of the staff members specialized in OSINT. This specialist staff conducts open source investigations based on the requests of other teams. The depth of the investigation (including with automated OSINT tools) depends on the question of the team or the intelligence needs.

The specialist department of the MIVD used several acquired tools for automated OSINT during the investigation period. These were purchased some two to three years ago. The investigation using tools for automated OSINT is often conducted on the basis of a simple assignment, such as the question if a social media service profile can be linked to a target. Another possibility is an extensive investigation, where it can be reasonably foreseen in advance that a 'more or less complete view' of certain aspects of an individual's private life will be obtained. That requires the use of the investigatory power for the systematic collection of personal data from publicly accessible information sources (Section 38 of the ISS Act 2017).

Finally, almost all MIVD teams have appointed staff members to conduct open source investigations to a certain depth, but without using tools for automated OSINT.

2.7 Differences between automated OSINT at the AIVD and the MIVD

The CTIVD's general observation is that the OSINT department at the MIVD is more advanced in automated OSINT than at the AIVD. By way of illustration: the investigatory power under Section 38 of the ISS Act 2017 was deployed 73 times by the MIVD and only once by the AIVD for the systematic collection of data about an individual from publicly accessible information sources using an automated OSINT tool. A single deployment of the investigatory power under Section 38 of the ISS Act 2017 can also be directed at a limited and specified number of individuals.

The main difference in the number of requests for systematic open source investigation using tools for automated OSINT (73 by the MIVD compared with once by the AIVD) can be explained by the fact that the use of tools for automated OSINT at the MIVD forms part of the standard intelligence process of its OSINT department and that the MIVD has more experience in their deployment. At the AIVD, the tools have been in a test phase for a long(er) period of time, with only a very limited number of staff having access to the automated OSINT tools.

3. Automated OSINT: Review of the legal framework

In this section, the CTIVD reviews to what extent the AIVD and the MIVD have acquired a sufficient understanding of the operation of the automated OSINT tools and the origin and the nature of the underlying sources with a view to complying with the data processing provisions.

The CTIVD conducts its review specifically on the following legal provisions:

- The provisions regarding data processing in the ISS Act 2017;
- The heads of service ensure the secrecy of the qualifying data, the qualifying sources from which the data is derived and the security of the individuals with whose cooperation the data is collected (Section 23 of the ISS Act 2017);
- The heads of the services ensure that the technical, staffing and organizational measures relating to data processing comply with the provisions under the law (Section 24 of the ISS Act 2017).

Prior to the development or deployment of a tool for automated OSINT, the services must ask themselves how the future data processing relates to the ISS Act 2017. That is because the functioning of a tool and the origin and nature of the sources consulted in the process will have an impact on the severity of the breach of fundamental rights of the data subjects and thus the proportionality assessment to be conducted. This information is also important in order to assess beforehand if the investigation might be systematic. Furthermore, based on that information, the head of the service is able to ensure lawful data processing and take measures if necessary (Section 24 of the ISS Act 2017).

3.1 General provisions regarding data processing

The data processing using tools for automated OSINT by the services must comply with the general provisions regarding data processing. The requirements under Section 18 of the ISS Act 2017 are key in that respect. Section 18 of the ISS Act 2017 regulates the general principles for data processing. Section 18(1) of the ISS Act 2017 stipulates that data may only be processed for a certain purpose and only in so far as necessary for the proper execution of the ISS Act 2017 (or the Security Screening Act). Section 18 (2) of the ISS Act 2017 stipulates that data should be processed fairly, carefully and in accordance with the law. The former means that an assessment of necessity and proportionality must take place. Article 18(3) of the ISS Act 2017 stipulates that the data processed in the context of the services' tasks must be labelled with the degree of reliability or a reference to the document or source from which the data is derived.

For the assessment of necessity, proportionality and reliability as well as accuracy of the data to be conducted, it must be clear – before a tool for automated OSINT is deployed – how the tool functions and what the underlying data sources are.

3.2 Proper and careful data processing

Proper and careful data processing also means checking what the purpose of the processing is, whether the data processing is necessary to achieve that purpose and if the data processing is proportional, in light of the breach of the fundamental rights of the data subject or subjects. Before using the tools for automated OSINT no decisions were made about proportionality. The services checked which operational value the tools with underlying sources had for them and in doing so assessed the requirements of purpose and necessity, but they failed to adequately check the impact on the fundamental rights of the data subjects.

Prior knowledge of the working (functionalities) and underlying sources of the tools is necessary to be able to conduct the proportionality assessment. Obviously, staff can (while copying data) decide between the interests at stake in processing the data for the relevant intelligence investigation and the severity of the breach of the fundamental rights of the data subject. However, this is not enough. Based on a prior assessment and later on the experiential learning an estimate can be made, for example, whether sensitive information will be processed by the deployment of certain sources in tools for automated OSINT. However, sensitive data may only be processed in addition to the processing of other data and only in so far as this is unavoidable for the purpose of data processing (Section 19 of the ISS Act 2017). In some cases, the decision can be taken not to use certain functionalities of the tool or not to copy certain results. For example, if that could lead to a systematic use as a consequence of processing large amounts of location data.

The reason that a separate assessment is necessary, prior to the tool's deployment, is that the functionality of a tool and the nature of the sources could involve a different impact on the fundamental rights, in particular the right to privacy. The interference with fundamental rights of data subjects is minimal when, for example, data is copied from news articles and data on the publicly accessible parts of social media services.¹⁹ Those are the sources referred to in the explanatory memorandum to the ISS Act 2017. However, the tools for automated OSINT used by the AIVD and the MIVD can also contain location data and leaked data sets, as explained above. It is clear, including from European case law on the processing of location data by the intelligence and security services, that the processing of location data constitutes a more severe breach of the right to protection of personal data and privacy.²⁰ The processing of data from leaked data sets constitutes a greater breach of fundamental rights than, for example, the processing of data from publicly accessible news articles.²¹ In the Netherlands, making non-public data available and trading stolen data is criminalised in the Netherlands since 1 March 2019. The criminalization is indicative of how copying data from leaked data sets is viewed in Dutch society.²² This has an impact on the proportionality assessment when processing that type of data.²³

A systematic approach *beforehand* can help to make a decision on the breach of fundamental rights that could occur when data is processed using the tool and can help to decide, in the context of the duty of care under Section 24 of the ISS Act 2017, which technical, staffing or organizational measures should be taken to ensure lawful data processing. It is conceivable that an organizational measure, for example, means that only the staff of a certain department or with a certain job description are allowed to use the tools for automated OSINT. Sections 2.5 and 2.6 show that the organizational measure referred to above has now been taken by both services.

¹⁹ *Parliamentary Documents II* 2016/17, 34588, no. 3, pp. 39 and 55-56.

²⁰ See for example ECHR 8 February 2018, 31446/12, ECLI:CE:ECHR:2018:0208JUD00314412 (*Ben Faiza/France*) and CJEU 6 October 2020, C-511/18 and C-520/18, ECLI:EU:C:2020:791 (*La Quadrature du Net e.a./Premier ministre e.a.*) and CJEU 6 October 2020, C-623/17, ECLI:EU:C:2020:790 (*Privacy International/Secretary of State for Foreign and Commonwealth Affairs et al.*)

²¹ See also report no. 55 (2018) and the Assessment Framework on bulk data sets offered on the internet by third parties.

²² From 1 March 2019, the Computer Crime Act III entered into force (*Bulletin of Acts and Decrees* 2019, 67). Section 138c of the Penal Code makes the copying or passing on of data from a non-public source an offence and Section 139g of the Penal Code makes trading of stolen data an offence.

²³ The Assessment Framework to report no. 55 (2018) on bulk data sets offered on the internet by third parties also observes that it is relevant to the severity of the privacy breach if a data set ended up in the public domain due to a criminal offence, such as hacking (computer intrusion (Section 138ab of the Penal Code)).

3.3 Reliability and accuracy of data

The data processed in the context of the execution with tools for automated OSINT must be labelled with the degree of reliability or a reference to the source from which the data is derived (Section 18(3) of the ISS Act 2017). In addition, the head of the service is responsible for taking measures to promote the accuracy and completeness of the data, and of the quality of the data processing (Section 24(2) part (a) of the ISS Act 2017).

The origin and accuracy of the data in the underlying data sources that can be consulted with the tools are not always clear. The origin of data is harder to verify, for instance, if that data consists of data sets compiled by the provider or of commercially offered location data. At the same time, tools for automated OSINT can contribute to promoting the accuracy and completeness of the data because it can be used to consult multiple data sources. In that way, a result from a single source can be verified by the result of another source, before the data is included in a report.

Investigation by the CTIVD showed that no adequate prior assessment had been made of the reliability and accuracy of the data prior to deployment of the automated OSINT tools. When the accuracy of the data is in doubt, the results from a search with automated OSINT tool are validated by AIVD and MIVD staff, using the results from another tool or data source. After deployment of the tool, staff will also record in their report what the underlying sources are on which the copied result is based. Although that is constructive, it is not enough for proper compliance with the general provisions on data processing.

The services must scrutinize the functioning (functionalities) of a tool and its underlying sources before using that tool. A systematic approach beforehand makes it possible to assess the reliability of certain sources in advance and – in the context of the duty of care under Section 24 of the ISS Act 2017 – to decide which staffing, organizational or technical measures need to be taken to bring the data processing into compliance with the law. One conceivable example of a staffing measure is to issue authorized staff with instructions explaining how to deal with certain underlying sources for an automated OSINT tool.

3.4 Automated data analysis

Data processing using tools for automated OSINT is fully subject to Section 60 of the ISS Act 2017. Simply put, it means that automated data analysis may not take place on the basis of automated decision making.²⁴ When data is processed using automated OSINT tools, it constitutes automated data analysis because it includes a file comparison based on a query with data on the underlying sources.

The services' policy and procedure ensure that no measures are taken against an individual solely based on a query in the tool. The results of a query in the tool or tools for automated OSINT are processed in a report drawn up by a staff member. It is only on the basis of this report, often in combination with other data, that measures – if any – against an individual may be taken. This procedure respects the prohibition on automated decision making in Section 60 of the ISS Act 2017.

²⁴ Section 60(3) of the ISS Act 2017.

3.5 Compliance with the duty of care for source and identity of staff

Before an automated OSINT tool was deployed, the AIVD and the MIVD looked expressly at operational risks (such as unauthorized individuals becoming aware of operations) and the security of the services' staff. The risk of the identity of individuals or organizations who are the subject of the investigation becoming known were also addressed and issued with a risk assessment.

The investigation established that the AIVD and the MIVD paid sufficient attention to protecting the identity of individuals or organizations under investigation by the services, the identity of staff conducting the queries and the data sources used in automated OSINT.

The head of the service ensures the secrecy of the qualifying data and the security of the individuals with whose cooperation the data is collected (Section 23 of the ISS Act 2017). The CTIVD established that the heads of the services duly observed the duty of care for the secrecy and security of individuals with automated OSINT.

3.6 Compliance with the duty of care regarding data processing

The heads of the services ensure that the technical, staffing and organizational measures relating to data processing comply with the provisions under the law (Section 24 of the ISS Act 2017). During the investigation period no prior decision or assessment was made relating to the privacy risks with automated OSINT.

The findings in section 2 on setting up a process for automated OSINT and the findings in this section on assessing compliance with the general provisions on data processing beforehand reveal that both services to some extent took more general staffing, technical and organizational measures to guarantee lawful data processing. The CTIVD pointed to the staffing and organizational measures such as the specifically designated staff members who conduct more in-depth OSINT investigations, with or without the use of automated OSINT tools. The findings from OSINT are documented and to some extent an assessment on reliability and accuracy of the data is made during the investigation.

Nonetheless, the CTIVD also established failings in the services' policy, procedures and work instructions as regards the organizational measures taken. That is negligent. A significant failing at the AIVD is caused by the flawed application of the criterion of a systematic approach in the context of automated OSINT. That criterion is important because the investigatory power in Section 38 of the ISS Act 2017 must be requested in writing and by stating reasons, in other words, authorization must be requested for the use, when it can be reasonably foreseen beforehand that by copying personal data from publicly accessible information sources a 'more or less complete view of certain aspects of an individual's private life will be obtained.'²⁵

The AIVD writes in its policy that using an automated OSINT tool to conduct a single query in the system of a target's name is never considered to be systematic. However, it is conceivable that a 'simple check' with the investigated tools will yield a large amount and variety of data about an individual, resulting in a more or less complete view of certain aspects of their private life. By then copying data on the target, for example from various social media services and data from different data sets, the investigation can in fact be systematic in nature. Based on the purpose of the investigation, knowledge of the tools and the underlying sources and experiential learning, an assessment must be made beforehand whether the query and subsequent copying of the data will be systematic.

²⁵ See also the Assessment Framework (Appendix I).

The MIVD has paid a great deal of attention to the criterion of systematic approach in its policy and work instructions for OSINT (more extensively than the AIVD). However, the CTIVD notes a significant failing in the policy and work instructions. The MIVD policy contains an exception to the criterion of a systematic approach by stating that copying 'business information' on 'fighters' will never be systematic. Data is considered to be business information when the purpose of the check is to obtain insight into the modus operandi, activities and business network of the fighter. However, there are no legal grounds for limiting the scope of the term systematic in that way. Indeed, it increases the likelihood of unlawful conduct in practice, because of a failure to use the authorization requirement for systematically collecting personal data from publicly accessible sources, for example. During the investigation by the CTIVD, the MIVD removed this exception from its policy and work instructions.

4. Conclusions and recommendations

The Intelligence and Security Services Act 2017 (ISS Act 2017) permits the AIVD and the MIVD to collect and process publicly accessible data, including personal data. These activities are also referred to as open source investigation or OSINT, which stands for 'open source intelligence'. The legislator does not consider OSINT to be an intrusive intelligence tool. It is regarded as a general investigatory power of the services, whereby a distinction is made between non-systematic and systematic deployment. Authorization is required for the systematic collection of personal data from information sources accessible to the general public.

When open source investigation is automated using specialized software or web applications, it is referred to as 'automated OSINT'. This investigation makes a distinction between the *tools* that are used and the *sources* (data sets) that may be accessed through these tools. The tools are in fact software equipped with search and network analysis functions which can query a wide variety of sources. These tools may come from commercial providers or be developed by the services themselves.

By conducting this review, the CTIVD aims to answer the following research question:

Do the AIVD and the MIVD have a sufficient understanding of the workings of the automated OSINT tools and the origin and the nature of the underlying sources with a view to complying with the data processing provisions?

In the present investigation, the CTIVD's lawfulness assessment focuses primarily on the OSINT tools and the data sets (i.e. the sources) that can be accessed using these tools. How these tools and sources are used in actual cases does not fall within the remit of the review. The CTIVD feels it is essential for the services to set up the process of automated OSINT in line with the ISS Act 2017. To do so, the services will need to fully understand how the tools function, such as the functionalities that the tools offer and the underlying sources. That knowledge will help them learn how the tools and sources relate to the legal requirements for data processing and the legal investigatory power to conduct systematic or other open source investigations.

Section 2 provides an outline of the different ways in which automated OSINT is conducted in general and more specifically within the AIVD and the MIVD. Section 3 lists the findings on compliance with legal requirements regarding data processing in the practice of automated OSINT at the AIVD and the MIVD.

Conclusions section 2

Section 2 shows that during the investigation period both the AIVD and the MIVD used multiple commercial tools to conduct automated OSINT. The MIVD uses these tools on a far greater scale than the AIVD. By way of illustration: during the investigation period from 1 July 2020 to 31 March 2021, the MIVD deployed the investigatory power to systematically collect personal data using a tool for automated OSINT a total of 73 times and the AIVD only once.

Tools for automated OSINT offer two major advantages over standard open source investigation using a web browser. The first of these is ease of use: a single search using an automated OSINT tool can query hundreds of sources simultaneously. The tool can then provide a visual representation of the results. The second major advantage to the services of using such tools is that they give access to sources based on user-friendly services provided by the tool's vendor on a commercial basis. One example of this is leaked data from users of social media services. Vendors can aggregate these data sets as a single searchable source (a 'composite data set'), which in some instances may contain billions of data points.

One example of commercial data that can be accessed through these tools is location data generated by ads shown to application users. Providers of commercial tools for OSINT can purchase advertisement data from data brokers and use their tool to make it available to clients, including intelligence and security services.

The disadvantage of a market product is that it is not always clear how the tool functions exactly and that only limited influence can be exercised on functionalities such as search options and logging. In certain cases, data sets cannot be purchased without the tool because that is part of the revenue model.

The volume, nature and types of personal data in these automated OSINT tools may lead to a more serious violation of fundamental rights, in particular the right to privacy, than consulting data from publicly accessible online information sources, such as publicly accessible social media data or data retrieved using a search engine.

From the explanatory memorandum of the ISS Act 2017, it can be concluded that the practices facilitated by automated OSINT were not taken into account by the legislator at the time and will continue to develop in the near future. OSINT undeniably goes well beyond investigative techniques such as checking telephone directories or using a search engine to access online data. The present investigation reflects the following reality: automated OSINT provides simultaneous searchable access to hundreds of sources of various origins, including location data or data from leaked data sets. The current practice of automated OSINT involves a more serious violation of privacy than was anticipated when the Act was drafted.

These findings lead to **recommendation 1**:

Given the nature, diversity, and volume of the data at issue, the Review Committee on the Intelligence and Security Services (CTIVD) recommends that the legislator creates a more foreseeable legal basis with sufficient safeguards governing the use of automated OSINT for both the tools themselves and the sources that can be accessed using these tools.

Conclusions section 3

Section 3 looks at the extent to which the services comply with the general provisions regarding data processing when processing data with tools for automated OSINT. The duty of proper and careful data processing in Section 18 of the ISS Act 2017 is key. Before a tool for automated OSINT is deployed, it is important that the functioning (functionalities) of the tool and the underlying data sources are scrutinized as best as possible. That knowledge can be used to check which data is processed and how. That information is necessary to conduct the necessity and proportionality assessments and check what the reliability and accuracy is of the data being processed by the tools. Sensitive data may only be processed in addition to the processing of other data and only in so far as this is unavoidable (Section 19 of the ISS Act 2017) (see sections 3.1-3.3).

The answer to the investigative question is that the AIVD's and the MIVD's understanding of the workings of the automated OSINT tools and the origin and nature of the sources that can be consulted using these tools is insufficient to ensure compliance with the data processing provisions of the ISS Act 2017. In practice the services do already take into sufficient account the operational and security aspects when using tools for automated OSINT (see section 3.5). In the context of the duty of care for lawful data processing the services should identify the procedures and the tools' underlying sources and take mitigating measures to prevent unlawful conduct in future (see section 3.6). The CTIVD notes that several improvements are needed before automated OSINT can be brought into compliance with the law.

This finding leads to **recommendation 2**:

When selecting and acquiring tools for automated OSINT (and thereby selecting the underlying sources), the AIVD and the MIVD should also aim to ensure lawful data processing. Preferably, the services should work together to develop a joint policy framework with accompanying work instructions.

In the interests of legal certainty, lawfulness and operational effectiveness on the part of the services (based on the continuity of lawful data processing at OSINT), the CTIVD will enter into a dialogue with the services in order to arrive at a workable temporary assessment framework which the services will then translate into policy, procedures, and work instructions. This temporary assessment framework should, among other things, address the establishment of a prior assessment in light of the data-processing provisions, the criterion of a systematic approach to open source investigation, keeping records and the handling of sources where the origin and accuracy of the data cannot be clearly established.

In conclusion

The use of OSINT is not exclusive to the domain of the intelligence and security services, but also applies elsewhere in the national security domain (for example, at the National Coordinator for Security and Counterterrorism) and beyond (including other government bodies). This review report notes that OSINT has continued to evolve over the years, allowing for the use of tools that simultaneously consult hundreds of sources and of which the results are displayed rapidly, clearly and in context. These underlying sources may include location data or leaked data. The processing of such data constitutes a violation of the fundamental rights of data subjects that goes further than OSINT using standard search engines or social media services.

The CTIVD therefore asks the Minister of the Interior and Kingdom Relations and the Minister of Defence to bring this report to the attention of other government bodies and, when forwarding the report, to ask Parliament to bring it to the attention of the House of Representatives' Standing Committee on Digital Affairs.

WWW

8.4854 963.8712

1010101

Oranjestraat 15, 2514 JB The Hague
P.O.Box 85556, 2508 CG The Hague

T T 070 315 58 20
E info@ctivd.nl | www.ctivd.nl