Int. Cybersecur. Law Rev. (2021) 2:317–337
https://doi.org/10.1365/s43439-021-00042-7

Open source intelligence

Introduction, legal, and ethical considerations

Isabelle Böhm · Samuel Lolagar

Received: 22 August 2021 / Accepted: 8 September 2021 / Published online: 24 November 2021

Abstract Open Source Intelligence (OSINT) has gained importance in more fields of application than just in intelligence agencies. This paper provides an overview of the fundamental methods used to conduct OSINT investigations and presents different use cases where OSINT techniques are applied. Different models of the information cycle applied to OSINT are addressed. Additionally, the terms data, information, and intelligence are explained and correlated with the intelligence cycle. A classification system for entities during OSINT investigations is introduced. By presenting the capabilities of modern search engines, techniques for research within social networks and for penetration tests, the fundamental methods used for information gathering are explained. Furthermore, possible countermeasures to protect one's privacy against the misuse of openly available information as well as the legal environment in Germany, and the ethical perspective are discussed.

Keywords Open Source Intelligence · OSINT · Privacy · Cybersecurity · Social Media

Isabelle Böhm · Samuel Lolagar (✉)
Hochschule Albstadt-Sigmaringen, Anton-Günther-Str. 51, 72488 Sigmaringen, Germany
E-Mail: info@OSINTgeek.de

Isabelle Böhm
E-Mail: isabelle@boehm.global

1 Introduction

1.1 Definition

The term Open Source Intelligence (OSINT) originally refers to a specific source of intelligence. In general, intelligence sources serve the purpose to produce raw data which can be further processed during the six steps of the intelligence cycle to gain insights (Office of the Director of National Intelligence 2011). **Open Source Intelligence** is defined as intelligence produced from publicly available sources that is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement (Office of the Director of National Intelligence 2011).

1.2 History

Its origins are credited to William Donovan (Colquhoun 2016). During World War II, he established the Office of Strategic Services which would later become the Central Intelligence Agency (CIA), the United States Foreign Intelligence Service. In this organization, an entire branch was tasked with the analysis of publicly available information and collected newspapers, journals, and even radio broadcasts from around the world bearing in mind the words of its founder that *"even a regimented press will again and again betray their nation's interests to a painstaking observer"* (Colquhoun 2016).

1.3 The Internet

The Internet is primarily a design concept for a resilient network architecture and includes all globally interconnected devices and the protocols used for communication. The Internet evolved from the ARPANET and the underlying idea was to create a decentralized network which is resilient and therefore able to ensure communication even after failure of central nodes (Abbate 1994). The most important protocol is the Internet Protocol (IP), which is responsible for routing and transport of packages between devices. The terms Internet and World Wide Web (WWW) are often used interchangeably, but this is not correct. The Internet is the fundamental technology and the WWW is a service built on top of it. There are other services that use the Internet e.g. email, but these services are not in scope of this paper. The WWW is often further divided into the **surface web** and the **deep web**. Most of the information is easily accessible using a web search engine, because the content is indexed by the search engine. This part of the Internet is called **surface web**. In contrast, the **deep web** can only be accessed by direct connection using the Unique Resource Locator (URL) or IP address, because its content is not indexed by a search engine (Thompson 2016). Some of these destinations require registration and sometimes also payment (Wikipedia Contributors 2021). Another part of the Internet is the so called **darknet** which consists of small peer-to-peer networks and requires certain software, configuration, and authorization to access (Wikipedia Contributors 2021). On top of the darknet the **dark web** is built. Here, sites, also referred to as *hidden services* are provided, which are only reachable by using the dark net (Thompson 2016). Thompson lists different darknet technologies – TOR, I2P, and FreeNet are examples for so-called overlay networks (Thompson 2016). Overlay networks use an existing network to create a new layer on top of it.

The rise of the Internet fundamentally changed the nature of public information. Focusing on newspapers and journals, most of them are now published in a digital version with frequently updated content. This results in a faster news cycle and easier access in contrast to earlier times. In addition to this well known, but digitally enhanced public sources, the Internet offers more options allowing every Internet user to share news, opinions, or knowledge without any extra editing and with very few restrictions. This is empowered by a variety of newly created public sources including personal websites, message boards, online encyclopedias, newsletters, blogs, or news groups. Companies and organizations also take the opportunity to publish information or provide online services like web search engines, social media platforms, trade directories, or dating websites.

These services are widely used and encourage users to voluntarily share a wide range of personal information linked to a pseudonym or even directly to the individual's full name. This emphasizes that the changes of public information are not limited to the sheer amount of information, but also concern the kind of information. Today's behaviour to voluntarily share personal information is in stark contrast to the reactions in the face of the 1983 German census where plans to survey some personal data were met with mass protest (Bundeszentrale für politische Bildung 2017).

Although some of these sources require additional efforts to access, all are still considered public as they are in general accessible compared to closed sources or classified documents. The lifespan of these sources rank from only a few years (cp. MySpace Stern Online 2019) to longer periods of over a decade (cp. LinkedIn W. Contributors 2021).

The resulting mass of information is made available by the Internet to a broad audience, not necessarily limited to the intelligence community. Motivated by different reasons, a variety of parties develops and utilizes best practices, tools, and techniques to collect, exploit, and disseminate this publicly available information to address their specific requirements.

2 OSINT Use Cases

In the last section, the statement was made that OSINT is not longer exclusively used by the intelligence community, rather by a number of other parties. This section illustrates this claim and shows a number of use cases observed in the last years where people with different backgrounds and motivations utilized different kind of public information.

Intelligence In 2015, a jihadist posted a selfie in front of an Islamic State bomb factory revealing the structure of the building. 23 hours later, the US military launched an attack destroying the building (Colquhoun 2016).

Journalism Bellingcat is a collective of researchers, investigators, and citizen journalists using open source and social media investigations to probe a variety of different subjects with impressive results. These include the identification of Russian intelligence officers as the key suspects in the Malaysian Airlines Flight 17 investigation as well as in the Skripal family poisoning. Moreover, they provided analysis of the chemical attack in Douma, Syria and of drone usage by non-state actors in Syria and Iraq. They exposed a fake persona who had been widely cited in Ukrainian and anti-Putin Russian media as a Pentagon official and revealed the illegal shipping of precursors of the nerve agent sarin to Syria by Belgian companies (The Bellingcat Collective 0000).

Recruiting A study from 2012 shows that already 13% of Dax and MDax companies initiated pre-employment screenings and background checks including online research during the hiring phase of a potential employee (Deloitte 2012).

Law Enforcement Agencies Starting in 2016, the German Police University conducted a three-year research study to evaluate how OSINT might provide relevant information for daily law-enforcement tasks and, thereby, reduce the risk for police forces as well as the general public (Epple and Ludewig 0000). Epple and Ludewig conclude that *"by implementing Open Source Intelligence in police dispatch centers, mission-relevant information can be obtained and that OSINT is a suitable instrument to ensure more professional mission accomplishment, better protection of the population and better self-protection of police officers."* (Epple and Ludewig 2020). Press releases suggest that the German police are now commencing to hire personnel tasked with OSINT investigations (Niedersachsen 2019). Foreign law enforcement agencies utilized OSINT in a number of cases including the investigations of a terrorist attack and an armed robbery as well as for a missing person search, a manhunt for a sex offender, and preparing an undercover persona (Ramwell et al. 2016).

Penetration Testing An important step while preparing a penetration test is the collection of data (Rieger et al. 2021). Depending on the defined goals, this can be facilitated by the use of tools and services exploiting public information for example the enumeration of a website's subdomains using tools like Sublist3r (Aboul-Ela 0000) or specialized search engines like Spyse (Spyse 0000) or Shodan (Shodan 0000).

Social Engineering and Human Intelligence Kevin Mitnick describes the first step of every social engineering attack as the collection and evaluation of information from available public sources (Mitnick and Simon 2002). Lekati even describes the combination of OSINT, Social Media Intelligence (SOCMINT), and Human Intelligence (HUMINT); *"OSINT and SOCMINT can be used as supporting disciplines when an investigator's or intelligence professional's ultimate goal is to be able to effectively interact with a suspect and either infiltrate a group, recruit the target, draw a confession or conduct other primarily HUMINT-related activities"* (Lekati and Lolagar 2021). OSINT and SOCMINT have large overlaps; SOCMINT emerged as a sub-discipline of OSINT and continues to be widely subsumed under Open Source Intelligence.

Public Tracing The law enforcement agency Europol established the project Stop Child Abuse in which the public is queried to identify image details using pieces extracted from sexually explicit material involving minors (Europol 0000).

Missing Person And Rescue Search After the mysterious disappearance of the Turing Award Winner Jim Gray during a boat trip in 2007 (May 2007), an unprecedented civilian search-and-rescue exercise was initiated including a crowd-sourced, automated analysis of satellite images and aerial views of the area where he disappeared (Hellerstein and Tennenhouse 2011).

Civil Protection A very similar application can be found within civil protection units. *"Already after the devastating earthquake in Haiti in 2010, [Volunteer & Technical Communities] V&TCs were founded with the aim of processing and providing publicly available data to emergency forces and the population"* (Fathi 2021). In 2016 the first German Virtual Operations Support Team (VOST) was founded by the Federal Agency for Technical Relief (THW) in cooperation and with support by researchers from the University of Wuppertal. *"In operational situations, the 'Digital Situation Exploration' group obtains information from social media and other public sources using Big Data evaluation methods in order to process and present it in a user-friendly manner. This includes the identification of mis- and disinformation, but also the verification and geo-localisation of information that is relevant to the situation"* (Fathi 2021).

Cyber Risk Management The monitoring and analysis of public sources can support an efficient risk assessment. This can be backed by tools tailored to the specific requirements of the respective organization (Revell et al. 2016). During risk assessments, often services to estimate the current level of security are used. Service provider like BlackKite *"[use] Open-Source Intelligence (OSINT) and non-intrusive cyber scans to identify potential security risks"* (Blackkite 0000).

Preparation of a criminal act The socialite Kim Kardashian made headlines in 2016 when she was robbed at gun point in her hotel room in Paris and jewellery worth EUR six million was stolen (BBC 2016). When one of the suspects was caught, he confessed to the police that he and his accomplices analysed Instagram posts and other Internet sources in combination with information from someone close to Kardashian to plan and commit the crime (Bryant 2017).

To google A basic tool to access information available in surface web is the usage of a web search engine with Google being the most popular one. Conducting an online search has became so common that there is a verb to describe this action, to google. The verb was included in the standard dictionary of the German language, the Duden, in 2004 (The Duden 0000). An addition to the Duden is only done after a selection process where the candidate word has to prove that it is used regularly over a longer period of time in different contexts (The Duden 0000).

The first use case shows an impressive example how the intelligence community used information gained by OSINT, but the rest of the use cases demonstrated that OSINT is applicable to a broad area. The common use of the verb ’*to google*’ highlights that the exploitation of public sources does not rely on a certain skill set. Experienced investigators may produce elaborated results beneficial to a broad number of different subjects, but also untrained people are able to derive certain information. The versatility of possible applications combined with its availability to (almost) everyone makes OSINT a powerful method which will be further analyzed in the work at hand.

3 Contribution

The previous chapter presented an overview of OSINT. After a general introduction to the topic, its easy and versatile usage with a selection of use cases was displayed. The next chapter formalizes the methodology in which this paper’s contribution is threefold. Sect. 4.1 describes the process for the acquisition of information using OSINT. Sect. 4.2 characterizes the output derived from open sources, and introduces a classification. Finally, Sect. 4.3 presents a selection of tools contributing to OSINT investigations. This is followed by a discussion about possible countermeasures and the legal environment in Germany in Sect. 5.2 as well as a brief consideration of the ethical dimension in Sect. 5.3. The paper concludes with a summary in Chapt. 6.

4 Open Source Intelligence Methodology

This chapter presents the theoretical framework of an Open Source Intelligence operation. The process how to undertake an OSINT investigation is outlined, the terms **data**, **information**, and **intelligence** are clarified, and selected tools and techniques are presented.

4.1 Modelling the Process of an OSINT Investigation

Different models to formalize the process of an OSINT investigation exist. In order to transform raw data into actionable intelligence, the intelligence community derived a model called **intelligence cycle** (Office of the Director of National Intelligence 2011). It is applied to all sources of intelligence and, in particular, to OSINT. This model has been adopted by Gibson (Gibson 2016) and with some adjustments also by Hassan (Hassan and Hijazi 2018). Bazzell presents a practical interpretation which is used as a mandatory training manual by U.S. government agencies (Bazzell 2021). Other works emphasize information gathering and analysis and, therefore, introduce models focusing on these tasks. This applies to the comprehensive three-step model derived by Pastor, et al. (Pastor-Galindo et al. 2016), as well to the model of Tabatabaei, et al. (Tabatabaei and Wells 2016).

The intelligence cycle is visualized in Fig. 1. It contains six steps labeled **Direction**, **Collection**, **Processing**, **Analysis**, **Dissemination**, and **Feedback** which are described in detail in the rest of the section.

Direction This phase is dedicated to planning and preparation before the actual investigation initiates. Gibson summarizes this step as *"identification of intelligence required"* (Gibson 2016). More details are offered by Bazzell in Bazzell (2021). He describes a more practical approach which is detailed in the following. As Bazzell addresses OSINT operations related to human activities, his explanations are enhanced for cases in which IT-systems are targeted. An OSINT investigation is most likely to begin with a specific request or a mission assignment by a client. Bazzell lists the following examples as typical tasks: threat assessment of individuals or events, target profiles for individuals or organizations, account attribution, or subscriber identification. IT-Security related tasks might focus on threat assessment for a system or digital footprint of an organization. Bazzell recommends clarifying the given task and verify provided identifiers like real names, user names, or email addresses. If the target is an IT-system, a common identifier is the domain name, but other identifiers like organisation or company name as well as email addresses are possible. Further, he also suggests applying a technique called **triage**. This is an assessment to derive a plan that is likely to provide the best possible outcome for the future investigation. Moreover, he advises to prepare the technical environment with an instance of a virtual machine dedicated to this specific investigation equipped with applications and tools relevant for OSINT. He also offers instruction on how to set up such a virtual machine addressing the requirements related to a search for human activities.

Collection This phase focuses on the collection of data and is described as gathering data. The idea is to systematically search public data using the known identifiers and link the findings to produce results. Bazzell gives the most detailed account about this phase by structuring it in three steps (Bazzell 2021). First, he suggests invoking specialized search engines, websites, and services which might require a payment or a fee. If the responsible investigator is associated with a law enforcement agency, this includes their respective closed source information systems. Second, the investigation continues with an initial web search of the identifiers followed by the utilization of selected OSINT application, tools, and techniques depending on the target of the investigation. In order to structure this utilization, he derives a workflow for each of the identifiers email address, user name, real name, telephone number, domain name, and location. Each workflow starts with a given identifier and proposes different paths including specific tools. This results in new pieces of information which can be further exploited to gain more insights. For example, a given identifier user name is potentially helpful to identify real name, email address, or a social network profile. The workflow includes several approaches how to proceed. One approach is a manual check for all social networks for the given user name, thereby potentially identifying the real name. Another path described in the workflow is the guessing of the email address based on the provided information. A third path takes the user name as input into a set of tools provided by Bazzell. In addition, the input

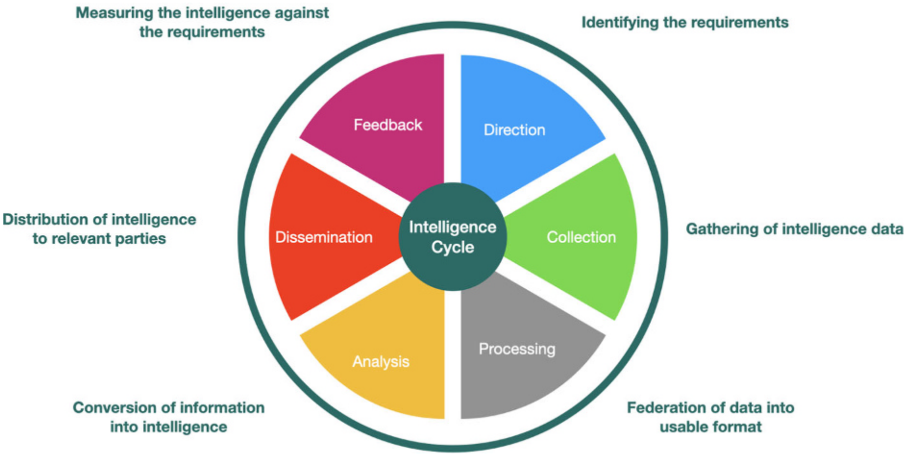
Measuring the intelligence against the requirements
Identifying the requirements
Feedback
Direction
Distribution of intelligence to relevant parties
Dissemination
Intelligence Cycle
Collection
Gathering of intelligence data
Analysis
Processing
Conversion of information into intelligence
Federation of data into usable format

Fig. 1 Visualization of the intelligence cycle as described in Office of the Director of National Intelligence (2011); Gibson (2016)

is processed by standard and specialized web search engines and further enriched with information from compromised databases. All these workflows, however, are in most cases tailored to a search located in the USA. In the third and final step of the collection phase, all findings are captured.

Processing This phase serves different purposes depending on the underlying model. The intelligence community and Gibson describe this step as the transformation of the collected data into information (Office of the Director of National Intelligence 2011; Gibson 2016). This implies translating, decrypting, or converting the data in an useful and understandable format. The models utilized by Bazzell, Pastor et al. do not specifically mention this step (Bazzell 2021; Pastor-Galindo et al. 2016) while others repurpose it for data enrichment (Tabatabaei and Wells 2016) or data verification (Hassan and Hijazi 2018). Although not explicitly indicated as a dedicated phase in the models of Bazzell, Hassan, Pastor et al., or Tabatabeai et al., the task of data transformation into information is not neglected. They include the required effort in the Analysis phase.

Analysis This phase converts information into intelligence as described by Gibson (Gibson 2016). This includes the integration, evaluation, and analysis of the gained information to produce a result meeting the requirements (Office of the Director of National Intelligence 2011). Bazzell points out that this step aims to understand how information is connected and how to represent these connections. Therefore, he advises to use a link analysis tool in order to visualize the results of the investigation (Bazzell 2021). This phase is split up by Pastor to emphasize that the analyzed information can be subjected to additional data mining or artificial intelligence techniques in a dedicated knowledge extraction phase (Pastor-Galindo et al. 2016).

Dissemination This phase distributes the results of the investigation to the client (Gibson 2016). This might be provided in the form of a written report (Bazzell 2021).

Feedback This phase concludes the investigation. While the U.S. National Intelligence and Gibson include the evaluation of feedback to improve their processes (Office of the Director of National Intelligence 2011; Gibson 2016), Bazzell finishes an investigation with the archiving of the results and a cleanup process (Bazzell 2021). The rest of the models renounce this phase (Pastor-Galindo et al. 2016; Tabatabaei and Wells 2016; Hassan and Hijazi 2018).

4.2 Data, Information, and Intelligence

The last section showed the process of an OSINT investigation not only as the acquisition of data but as the transformation from collected data into information and, finally, to intelligence. This section clarifies the terms **data**, **information**, and **intelligence**. Next, a proposed classification of information collected and compiled during an investigation is presented.

The differentiation between data, information, and intelligence derives from the NATO Open Source Intelligence Handbook and is frequently included in the discussion about the methodology of OSINT investigations, for example by Gibson or Hassan (Gibson 2016; Hassan and Hijazi 2018). This transformation process from data to intelligence is visualized in Fig. 2.

Data describes the output achieved during the collection phase in Fig. 1. It is considered as a set of facts without any explanation or analysis (Hassan and Hijazi 2018). The acquired data can be classified with respect to its format (Gibson 2016). It can be differentiated between structured data, semi-structured data, and unstructured data. Structured data is organized by an underlying data model and results in a standard format which is easy to process automatically. An important example for structured data are SQL databases (Gibson 2016). Although not as rigidly organized as structured data, semi-structured data also contain some structural elements. Examples include JavaScript Object Notation (JSON), Extensible Markup Language (XML), and Hypertext Markup Language (HTML) documents (Gibson 2016). Websites, reports as well as images, audio, or video are considered unstructured data (Gibson 2016). Fig. 2 depicts this output as **open source data**.

Contingent on the underlying model, **information** is either the output of the collection or the processing phase in Fig. 1. It is produced by processing the collected data. Depending on the nature of the data, processing includes translation, decryption, or format conversion, additionally filtering, correlating, classifying, clustering, and interpreting the given data. Fig. 2 refers to this output as **open source information**.

Compiling information to address a specific query results in **intelligence** (Gibson 2016). It is the result of the integration, evaluation, and analysis of information during the analysis phase in Fig. 1. Fig. 2 denotes it as **open source intelligence**.

Until this point, the phases of the intelligence cycle in Fig. 1 map to the different outputs in Fig. 2. However, Fig. 2 depicts a fourth and additional type of output which is not covered within the intelligence cycle. According to the NATO it is called **validated open source intelligence** (Gibson 2016). It is described as open source intelligence which *"a high degree of certainty can be attributed"* (Gibson 2016). This demands verification and validation of the derived open source intelligence. This can be potentially done using other intelligence sources.

The transformation from data to information can result in an abundance of different kinds of information. Therefore, a classification to structure which kind of information can be expected from an OSINT investigation is proposed in this work. Although OSINT is a very active field and therefore a classification might need constant adaptions, this paper presents a base point for further discussions.

Information is classified with respect to the intended target of an OSINT investigation. Seven types of addressed entities were identified during this research. Acquired information can relate to:

Single Person Findings range from information about real life – like full name, address, employment, or financial information; to the online persona with user name, email address, or social media presence.

Group of People In particular, criminal investigations often do not only focus on a single person, but on a set of people to understand their personal relationships and interactions.

Organisation In this paper, an organization is understood as an entity with a defined and articulated purpose that distinguishes it from a person or group of people. Examples are companies, institutions, associations, or even countries. The number of interesting information includes details about business deals, strategic planning, customer relations, employees, customers.

Computer System This summarizes all information related to IT-systems. It contains information about domain names and existing subdomains, used software and their respective version, as well as open ports.

Event Observing interactions between people might lead to information about an event happening online or in real life with details about date, location, and participants.

Location This includes details about a physical address or a set of coordinates.

Object This covers every target not addressed in one of the above classes. It includes images and videos as well as their content.

This classification is not distinct, meaning that a target might fit in more than one class. For example, the information "Woodstock" can be classified as group of people, event, or location. Its classification depends on the context of the initial query which also influences how the information is further processed.

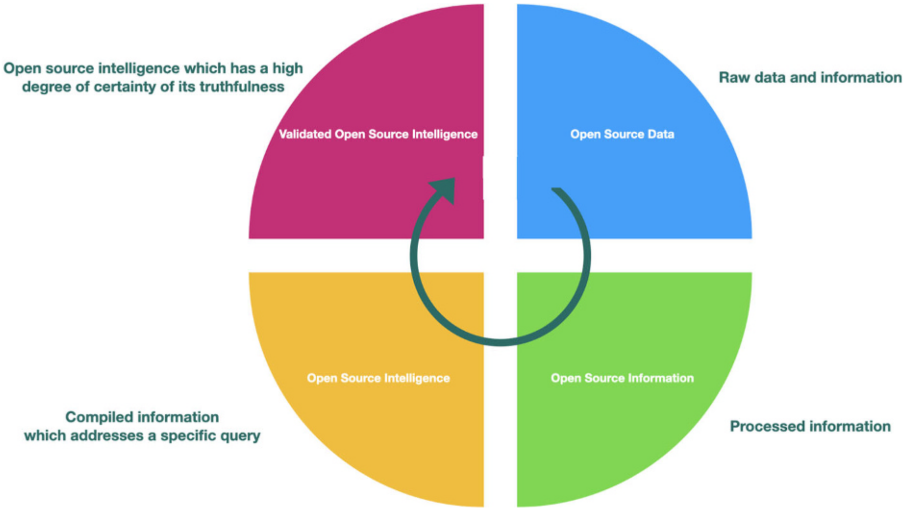
Open source intelligence which has a high degree of certainty of its truthfulness
Validated Open Source Intelligence
Open Source Data
Raw data and information
Open Source Intelligence
Open Source Information
Compiled information which addresses a specific query
Processed information

Fig. 2 Visualization of the data processing during an OSINT investigation as described in Gibson (2016)

4.3 Selected Tools and Techniques

Even if the intelligence cycle is initiated with a precise query as input, the response to this query relies on collecting, processing, and analyzing massive amounts of public data. This is predicated on at least the partial automation of certain tasks. In particular, the collection phase can be facilitated by the utilization of different tools and techniques.

A complete overview on tools is difficult to provide in light of the fact that many specialized tools are utilized. In addition, the landscape of external tools is extremely active and subject to change. One reason for changes is the revocation of the tools by their developers as observed in June 2019 when Bazzell withdrew his set of popular interactive online tools (Bazzell 2021) or the disappearance of the meta-crawler website searx.me. Another aspect is related to the dynamic nature of social networks. For example, Facebook and Instagram are known to actively undermine the usage of OSINT related tools and techniques. Therefore, they block respective web services, regularly change their source code, and restrict capacities exploited by the OSINT community (Bazzell 2021). For example, Instagram includes special character encoding in the source code of their website to make it difficult to directly extract URLs.

Nevertheless, the OSINT community has shown some resilience in the face of these challenges and constantly adapts or renews its approaches. Given the aforementioned aspects, a selection of tools and techniques only represents a limited snapshot and might be obsolete soon.

Despite all these difficulties, different approaches to overview tools and techniques exist. THE OSINT FRAMEWORK is the most developed ressource (Nordine 0000). Its goal is the easy identification of OSINT tools suitable for the search based on specific identifiers. It is organized as a tree structure with identifiers as the roots and candidate tools as the leaves. In addition to this excellent classification, Michael Bazzell's selection of tools with respect to their usability should be highlighted. The tools are provided through the setup of a virtual machine according to the instructions found in Bazzell (2021). These tools are well documented, regularly updated, and maintained.

As the remainder of this section can only offer a limited overview, the focus is on the most promising starting points for an OSINT investigation of human activities as well as of computer systems. First, web search engines are discussed as a general basis followed by social network searches addressing the search for human activities and continue with information-gathering focusing on IT-systems.

4.4 Web Search Engines

Tarakeswar et al. describe four types of search engines (Tarakeswar and Kavitha 2011): Crawler-based search engines, human-powered directories, hybrid search engines, and meta search engines. The aforementioned search engines are considered general crawler-based search engines. They perform three operations: crawling, indexing, and searching. Crawling is the process of finding and reading the content of a website. At this point, a crawled website will not be listed in the search results. Therefore, indexing is necessary first. During indexing the information on the website is extracted and stored within the search engines index. By performing a search using a search engine, the index is queried and returns results from its index pointing to the crawled website.

As described in Sect. 4.1, the collection phase invokes a web search based on the provided information, the so-called **identifiers**. A web search relies on a web search engine which crawls the web, indexes the findings systematically, and supports a search over the results. Well known is the classic textual search where the input of one or more words initiates the search and the output is presented in the form of links to websites containing the input. In addition to the aforementioned example Google (Google 0000), there are several competitors offering similar services like Bing (Bing 0000), Yandex (Yandex 0000), Baidu (Baidu 0000), DuckDuckGo (Duckduckgo 0000), or StartPage (Startpage 0000). These services are a good starting point for identifiers like people's names, organisations or public events.

Compared to the simple textual search, a careful formulation of the search query might improve the results significantly. Therefore, the input to the well-known user interface is enriched with extended search parameters combined with Boolean expressions. **Extended search parameters** are special characters and commands which extend the capability of textual search significantly. Examples include the usage of quotation marks requesting an exact match, the term **intext:** for a search in the body or document, the term **inurl:** for a term in the URL. Their utilization produces refined results. Consider the following example: a search is made for PDF documents containing the term OSINT. The input *OSINT, PDF* produces almost 834.000 results of mixed file types compared to the input *OSINT filetype:PDF* with only 42.000 PDF-files as result. This technique was first applied in the Google search engine and is therefore known as Google Hacking (Long 2005) or Dorking (Hassan and Hijazi 2018). However, it is also applicable to some extent for other search engines as well. Johnny Long reported on this approach since the early 2000's, for example in Long (2005). In particular, this technique is useful for personal-related information as well as for system-related information.

Given identifiers like user names, phone numbers, or email addresses as inputs, the results of the described web search engines might be limited. This is also true for the analysis of other potential relevant content such as images or locations. Specialized search engines are able to address such kinds of inputs. These include image search (Google Images 0000), as well as searches for user names (Knowem 0000), email addresses (Hunter.io 0000), locations (Google Earth 0000), or phone numbers (Das Telefonbuch 0000). Looking beyond the surface web to the dark web, there are also different web search engines available (Not Evil 0000).

Other search engines are tailored to output only specific results, e.g. real names (True People Search 0000), news (Google News 0000), scientific papers (dblp – Computer Science Bibliography 0000), patents (Europäisches Patentamt 0000), security vulnerabilities (Common Vulnerabilities and Exposures 0000), or Internet-connected assets (Shodan 0000). Bazzell suggests the utilization of the programmable search engine by Google which allows customized searching and filtering for OSINT investigations (Bazzell 2021).

4.5 Social Network Searches

Promising sources of information about an individual or a group of people are social networks like Facebook, Instagram, LinkedIn, Twitter, Pinterest, YouTube, or even PayPal (Lolagar 0000) where people share personal information and interact with families, friends, colleagues, or even strangers. This approach is of particular interest if the given identifier is a user name and might finally result in the identification of the real name. However, a real name as a given identifier can also be exploited as it is customary to use real names in some social networks like LinkedIn and, to some extent, Facebook. Findings in one social network can be included for searches in other social networks and the merge of information leads to even more detailed results.

Before an OSINT investigation can collect any information, it is often necessary to register to the respective social network with Twitter being a notable exception as tweets and timelines are accessible even without an account. For social networks requiring registration, this step might be sufficient to learn a lot (if not all) interesting facts about an account in case the target has a public profile. Detailed analysis is supported by automated download tools adapted to a respective social network, for example InstaLooter (Instalooter 0000) or Instaloader (InstaLoader 0000) focusing on Instagram or TweetBeaver (Tweetbeaver 0000) or exportdata (ExportData 0000) focusing on Twitter.

Apart from a manual examination of a target's profile, an important technique for the collection of information in social networks is the submission of search queries. The simplest approach utilizes the internal search functionalities provided by the social network itself. However, search features vary depending on the respective social network. For example, Facebook's internal search is simple compared to

Twitter which offers search operators comparable the extended search parameters for web search engines. Facing such limitations, there are attempts to introduce search options beyond the designed scope. This relies on the knowledge and manipulation of certain URLs which are usually automatically created and deployed producing results for search queries which cannot be submitted directly. This can be automated by a range of tools, for example Bazzell's Facebook tool (Bazzell 2021).

In addition, there is a number of applications, browser extensions, or web services specialized on different social networks which offer the extraction of certain information either as a preparation for URL manipulation (e.g. the extraction of a Facebook userID Facebook UserID LookUp 0000) or as stand-alone information (e.g. displaying biography changes in Twitter Twitter Biography Changes 0000).

Challenges arise as some social networks allow to apply strong privacy settings on the profile which prevents a detailed examination. Although this omits a substantial amount of information, it does not necessarily prevent information leakage. This is documented by a number of practitioners' tutorials allowing to derive information about private profiles without compromising the target's user account. For example, it is possible to identify public posts not directly shown on the private profile page (OSINT Curious 0000). Another option is the analysis of connected accounts which might be public and reveal information about the target.

Tools and techniques for searches on social networks are constantly evolving. As cited before, this is caused by frequent changes by the social networks themselves to prevent the exploitation of information in ways not originally intended. A notable example is the disappearance of Facebook's graph search. Initially introduced for general usage, it offered a powerful approach to derive information. This changed in 2014 and graph search functionality was only available using URL modifications. Finally, all tools and techniques relying on graph search stopped working in mid 2019 (Bazzell 2021). The assumption that Facebook is actively preventing data collection in non-intended ways is further fuelled by the block of web services offering automatic exploitation. Furthermore, also Facebook user accounts which invoke OSINT-related tools and techniques are regularly blocked (Bazzell 2021).

4.6 Information Gathering for Penetration Testing

Easily overseen, OSINT is a key technique to support information gathering for penetration testing or similar IT-related tasks. As for all OSINT-related investigations, the utilization of web search engines and in particular Google Hacking is a valuable starting point (Long 2005). Similar to the tailored tools available for the search of human activities, a range of tools dedicated for investigation on IT-systems exist.

Depending on the given identifier, a multitude of tools allows further inspections. A company's website can be evaluated with respect to applied technologies (BuiltWith 0000) as well as possible vulnerabilities (Common Vulnerabilities and Exposures 0000). Further information can be gained by identifying which websites share the same Google Analytics ID (Reverse analytics id 0000) or inspecting the website's history using an Internet archive (The Wayback Machine 0000).

Given a domain name as identifier, tools provide information about the ownership (WhoIs Online Service 0000) as well as about possible subdomains (Aboul-Ela

0000). Open ports can be identified by specialized web search engines supporting the search over pre-scanned results (Censys 0000). In addition, there are also tools and services performing active port scans available (Nmap 0000). This functionality is sometimes enriched with additional information about unpatched security vulnerabilities (Cyberscan 0000).

In order to collect more information about the system landscape of a company, another approach includes the target's advertisements and employee profiles on social networks to learn about declared skills to conclude about the used technologies and systems.

The collected information can be analyzed by dedicated analysis tools like spiderSilk (Spidersilk 0000) or the version of the utilized technologies can be matched against a database containing vulnerabilities for specific versions.

5 Discussion

The preceding chapters provided an overview about OSINT and presented a selection of tools and techniques. It became evident that the goal of every OSINT investigation is to gain intelligence about a target. However, this might contradict the target's interests. From the target's perspective, disclosure of information may be undesirable as it undermines privacy, or even harmful if the results are leveraged against the target in a criminal act. This raises the question of whether there are measures to prevent the exploitation of data.

The contribution of this chapter is twofold. On the one hand, it recommends countermeasures against OSINT to preserve privacy in the face of such intrusion and, on the other hand, discusses the legal environment in Germany regarding the use of open source intelligence.

5.1 Countermeasures

In the face of the unwanted exposure of personal data, it is a valid question what each one can do to minimize the attack surface. The answer adheres to the well-known principle to keep data private and to share as little information as possible. The less public data is available about an individual, the lower the risk of being exposed during an OSINT investigation. There are some recommendations for this situation available (Bore 0000), but depending on the personal circumstances they might be difficult to implement. Next, some useful best practices for individuals are further described.

The first recommendation is a detailed review of social network activities and active reduction of the amount of shared information. The most efficient strategy would be to avoid participation in social networks completely and close all existing accounts. However, some might argue that interactions on social networks are indispensable in their day-to-day life. In this case, it is recommended to adjust the privacy settings of all profiles. This might include the use of a pseudonym instead of the real name. In addition, even with strict privacy settings in place, it is wise to treat all shared information as publicly available and avoid exposing content unsuitable for a broad audience or compromising in the wrong hands. This also applies to relationships in social networks. As these contacts might also be targeted during an investigation, they should apply these recommendations as well.

Aside from the review of social network activities, another piece of advice is to avoid the publication of personal information like home address, phone number, or email address at all. In case a publication is necessary with respect to business reasons, a feasible approach is to set up separate contact information.

5.2 Legal Environment

5.2.1 GDPR

In 2016, the European Union introduced the General Data Protection Regulation (GDPR) (The European Parliament and Council of the European Union 2016). This regulation introduces protection measures for personal data of natural persons. For example, it requires notification if personal data is collected and consent in order to process personal data. Fines are imposed in case of violation. Focusing on OSINT investigations on individuals, the application of this regulation would severely restrict any data collection or processing. Overall, it has been stated that legal issues around OSINT investigations are uncertain in many areas (Lyle 2016). Block states in his recent article that *"[...] even though OSINT researchers by definition collect their data from publicly available sources, they still should comply with the GDPR where applicable"* (Block 2021).

Especially regarding the constant monitoring of large parts of the Internet, the GDPR states *"[a] data protection impact assessment referred to in paragraph 1 shall in particular be required in the case of [...] a systematic monitoring of a publicly accessible area on a large scale"* The European Parliament and Council of the European Union (2016, Art. 35 n. 3).

5.2.2 Authorities

Regarding authorities there are concerns that OSINT could be seen as interfering with fundamental rights. The Federal Constitutional Court is the highest legal body in Germany with its decisions being final and binding (The Federal Constitutional Court 0000). Relating to public data, there was a judgement issued on February 27, 2008 stating that *"if the state obtains knowledge of communication contents which are publicly accessible on the Internet, or if it participates in publicly accessible communication processes, in principle it does not encroach on fundamental rights"* BVerfG (2008, Nr. 6). This can be interpreted as a legitimization to conduct OSINT investigations. The court further specifies that *"[t]he state is not in principle denied the possibility to obtain publicly accessible information. This also applies if personal information can be collected by these means in an individual case"* BVerfG (2008, para. 308). The Federal Constitutional Court even states that if content is made available on the Internet and is addressed to all readers or at least to a not further limited group of individuals, this does not constitute an encroachment on the general right of personality when a state agency collects this communication content BVerfG

(2008, para. 308). The judgement lists specific examples: opening of a website, subscribing to a mailing list, and monitoring of open chat rooms BVerfG (2008, para. 308). However, the court sets a limitation and clarifies that *"[a]n encroachment on the right to informational self-determination can however apply if information obtained by viewing generally accessible contents is deliberately compiled, stored and where appropriate evaluated using further data, and a special danger emerges from this for the personality of the person concerned."* BVerfG (2008, para. 309). Nevertheless, the Federal Constitutional Court makes a clear statement that *"[...] pure Internet reconnaissance will not as a rule bring about an encroachment on fundamental rights"* BVerfG (2008, para. 311).

With the introduction of the GDPR in 2016, processing of personal data is further regulated. This also applies in the event that the investigation is conducted by a law enforcement agencies with the EU directive 2016/680 specifically addressing the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties.

In addition, the collection of public information by authorities, namely law enforcement agencies, is subject of controversial discussions. While targeted searches are covered by general authorization clauses, the use of broad-based OSINT searches (e.g. Big Data) is in dispute.

5.2.3 Terms of services

Another interesting aspect arises from the question of whether the data collected is really public. In the course of this work, public data was described as the opposite of classified data or closed sources. Thereby, this point of view regards almost all data accessible in the Internet as public. However, this description deserves a closer analysis. In particular, Sect. 1.1 implied that data acquired from social networks can be considered as public. However, access to the most common social networks like Facebook or Instagram is restricted by a registration process. This means that only registered members of this social network are eligible to access. Although this registration is free of charge, it requires to enter into a contract with the social network. Thereby, the user has to agree to the terms of service of the social network. In some cases, the application of OSINT tools and techniques violates this agreement. For example, it is forbidden to access Facebook or Instagram in an automated manner according to their terms of service or to create a Facebook account using a fake name. Moreover, it is known that Facebook actively blocks OSINT related services (Bazzell 2021).

5.3 Ethical Considerations

In terms of the ethical perspective, it is also necessary to discuss whether the ethical use of information depends on the information being made public intentionally. For instance, in Sect. 4.5 PayPal was mentioned – here it is possible to search through a list of PayPal users by just providing a few characters of the name. Often the corresponding profiles include additional information like the current town of residence. Most people provided their name believing that only a limited number of people will have access to this information. Consequently, the majority of users are not aware that some of their information can be considered publicly available.

Breach data is often made public, either to underline the demand for a ransom or to prove a successful attack. With this kind of data being publicly available and the high value of the data, OSINT investigators wonder if it can be exploited for an investigation. And on this question, opinions differ. For some, the use is unethical because the data was never intended for the public and was obtained through criminal acts. For others, it is irrelevant how it was obtained, and their own collection is enriched. However, for others, what the data is used for is relevant.

With the golden rule and the fact that once people learn about OSINT techniques, almost all change their online behavior, leads to the conclusion that not everything that is possible on a technical level, is ethical justifiable.

6 Conclusion

This paper has provided an introduction to Open Source Intelligence and presented several use cases. In doing so, it became obvious how versatile and powerful the techniques can be applied. The intelligence cycle, as well as the relevant terms data, information, and intelligence, have been introduced. Furthermore, a classification of entities was presented. An overview about tools and techniques for information gathering using search engines, social media, and regarding computer systems was provided. It became obvious that the OSINT landscape is constantly morphing. Basic measurements to protect private information from being exposed by OSINT investigations are available, but not always very practical. The legal environment was discussed. Overall, the utilization of public data is, in most cases, deemed lawful in Germany allowing OSINT investigations focusing on computer systems, objects, or locations. The situation for OSINT investigations on individuals is different as such investigations rely on the collection and processing of personal data. Such types of data are subject to restrictions on the European level which makes the legal status of OSINT focusing on natural persons at least by non-governmental entities uncertain and subject to further clarification. From an ethical point of view, it can be noted that not everything that is technically possible is also moral.

Funding Open Access funding enabled and organized by Projekt DEAL.

Conflict of interest The author Samuel Lolagar is member of the Virtual Operations Support Team (VOST) of the Federal Agency for Technical Relief (Technisches Hilfswerk, THW), which is led by Ramian Fathi. Furthermore, he has designed a training in cooperation with Christina Lekati. Both were cited in this paper, due to their individual contribution. The authors declare that they have no conflict of interest.

References

Abbate JE (1994) From ARPANET to INTERNET: A history of ARPA-sponsored computer networks. Ph.D. thesis, University of, Pennsylvania, pp 1966–1988

Aboul-Ela A Sublist3r. https://github.com/aboul3la/Sublist3r. Accessed 21 Mar 2021

Baidu. www.baidu.com. Accessed 21 Mar 2021

Bazzell M (2021) Open source intelligence techniques – resources for searching and analyzing Online information, 8th edn. (self-published)

BBC (2016) Kim Kardashian West robbed of millions by Paris gunmen. https://www.bbc.com/news/world-europe-37538453. Accessed 26 Feb 2021

Bing. www.bing.com. Accessed 21 Mar 2021

Blackkite. https://blackkite.com/platform/. Accessed 25 Feb 2021

Block L (2021) GDPR essentials for OSINT research. https://www.blockint.nl/methods/gdpr-essentials-for-osint-research/. Accessed 28. 07 2021

Bore J Do you know how to stay protected? https://circuit-magazine.com/onsit-do-you-know-how-to-stay-protected/. Accessed 16 Mar 2021

Bryant K (2017) Kim Kardashian's alleged robber confirms social media helped him plan heist. https://www.vanityfair.com/style/2017/01/kim-kardashian-paris-robbery-social-media-heist. Accessed 26 Feb 2021

BuiltWith. https://builtwith.com/. Accessed 21 Mar 2021

Bundeszentrale für politische Bildung. Vor 30 Jahren: Protest gegen Volkszählung. https://www.bpb.de/politik/hintergrund-aktuell/248750/volkszaehlung-1987-22-05-2017 (2017). Accessed 20 Mar 2021

BVerfG (2008) Judgment of the First Senate of 27 February 2008 – 1 BvR 370/07. https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/2008/02/rs20080227_1bvr037007en.html. Accessed 14 Mar 2021

Censys. https://censys.io/. Accessed 21 Mar 2021

Common Vulnerabilities and Exposures. cve.mitre.org. Accessed 21 Mar 2021

Colquhoun C (2016) A brief history of open source intelligence. https://www.bellingcat.com/resources/articles/2016/07/14/a-brief-history-of-open-source-intelligence/. Accessed 25 Feb 2021

Cyberscan. cyberscan.io. Accessed 21 Mar 2021

Das Telefonbuch. www.dastelefonbuch.de. Accessed 21 Mar 2021

dblp – Computer Science Bibliography. dblp.org. Accessed 21 Mar 2021

Deloitte (2012) Riskominimierung bei der Personalauswahl. https://www2.deloitte.com/content/-dam/Deloitte/de/Documents/finance/Studie-Risikominimierung-bei-der-Personalauswahl.pdf. Accessed 8 Mar 2021

Duckduckgo. duckduckgo.com. Accessed 21 Mar 2021

Epple G, Ludewig F (2020) Schriftenreihe der Deutschen Hochschule der Polizei vol 11

Epple G, Ludewig F Sentinel: Sicherheit im Einsatz durch Open-Source-Intelligence (OSINT) in Einsatzleitstellen. https://www.dhpol.de/departements/departement_II/FG_II.1/projekt-sentinel.php. Accessed 9 Mar 2021

Europäisches Patentamt. www.epo.org. Accessed 21 Mar 2021

Europol. Stop Child Abuse. https://www.europol.europa.eu/stopchildabuse. Accessed 8 Mar 2021

ExportData. https://www.exportdata.io/download-twitter-account-timeline. Accessed 21 Mar 2021

Facebook UserID LookUp. https://miniwebtool.com/de/facebook-user-id-lookup/. Accessed 21 Mar 2021

Fathi R (2021) Social media in disaster events – challenges and possible solutions in a highly networked society. https://www.buw-output.uni-wuppertal.de/en/output-ausgabe-022021/social-media-in-disaster-events-challenges-and-possible-solutions-in-a-highly-networked-society/. Accessed 25 Feb 2021

Gibson H, in *Open Source Intelligence Investigation – From Strategy to Implementation*, ed. by B. Akhgar, P.S. Bayerl, F. Sampson (Springer, 2016), pp. 69 – 93

Google Earth. www.google.com/intl/de/earth/. Accessed 21 Mar 2021

Google Images. images.google.com. Accessed 21 Mar 2021

Google News. news.google.com. Accessed 21 Mar 2021

Google. www.google.com. Accessed 21 Mar 2021

Hassan NA, Hijazi R (2018) Open source intelligence methods and tools – A practical guide to Online intelligence. Apress, Berkeley

Hellerstein JM, Tennenhouse DL (2011) Searching for Jim Gray. Commun ACM 54:77–87

Hunter.io. hunter.io. Accessed 21 Mar 2021

InstaLoader. https://instaloader.github.io/. Accessed 21 Mar 2021

Instalooter. https://github.com/althonos/InstaLooter?files=1. Accessed 21 Mar 2021

Knowem. knowem.com. Accessed 21 Mar 2021

Lekati C, Lolagar S (2021) Why for today's cyber investigations we need to combine intelligence disciplines. https://christina-lekati.medium.com/why-for-todays-cyber-investigations-we-need-to-combine-intelligence-disciplines-afca5363048c. Accessed 25 Feb 2021

Lolagar S Wie man fasst jede E-Mailadresse oder Handynummer identifiziert. https://www.youtube.com/watch?v=Bo8JWBqBThI. Accessed 21 Mar 2021

Long J (2005) Google hacking. MITP, Frechen

Lyle A, in *Open Source Intelligence Investigation – From Strategy to Implementation*, ed. by B. Akhgar, P.S. Bayerl, F. Sampson (Springer, 2016)

May M (2007) Vast search off coast for data wizard. https://www.sfgate.com/news/article/Vast-search-off-coast-for-data-wizard-2620302.php. Accessed 10 Mar 2021

Mitnick K, Simon W (2002) Die Kunst der Täuschung – Risikofaktor Mensch. MITP, Frechen

Nmap – Security Scanner. https://nmap.org/. Accessed 21 Mar 2021

Nordine J The osint framework. https://osintframework.com. Accessed 14 Mar 2021

Not Evil. http://hss3uro2hsxfogfq.onion. Accessed 21 Mar 2021

Office of the Director of National Intelligence—United States of America. (2011) U.S. National Intelligence—An Overview 2011. https://www.dni.gov/files/documents/IC_Consumers_Guide_2011.pdf. Accessed 25 Feb 2021

OSINT Curious (2020) 10 Minute Tip: How to find Facebook data when a profile is private. https://www.youtube.com/watch?v=NqzvuUXkv6c. Accessed 21. Mar 2021

P. Niedersachsen (2019) 13 Intel Officer als Verstärkungen für die niedersächsische Polizei. https://www.mi.niedersachsen.de/aktuelles/presse_informationen/13-intel-officer-als-verstaerkungen-fuer-die-niedersaechsische-polizei-174810.html. Accessed 9 Mar 2021

Pastor-Galindo J, Nespoli P, Marmol FG, Perez GM (2016) IEEE Access 4. https://doi.org/10.1109/ACCESS.2020.2965257

Ramwell S, Day T, Gibson H in *Open Source Intelligence Investigation – From Strategy to Implementation*, ed. by B. Akhgar, P.S. Bayerl, F. Sampson (Springer, 2016), pp. 197 – 211

Revell Q, Smith T, Stacey R, in *Open Source Intelligence Investigation – From Strategy to Implementation*, ed. by Akhgar B, Bayerl PS, Sampson F (Springer, 2016), pp. 153 –165

Reverse analytics id. https://dnslytics.com/reverse-analytics. Accessed 21 Mar 2021

Rieger M, Schlichtenberger D, Scheible T (2021) Modul 106: IT-Sicherheit und IT-Angriffe Modulhandbuch

Shodan. https://www.shodan.io. Accessed 15 Mar 2021

Spidersilk. https://spidersilk.com. Accessed 21 Mar 2021

Spyse. https://spyse.com. Accessed 15 Mar 2021

Startpage. startpage.com. Accessed 21 Mar 2021

Stern Online (2019) MySpace: Darum spielt das Netzwerk keine Rolle mehr. https://www.stern.de/kultur/myspace-darum-spielt-das-netzwerk-keine-rolle-mehr-8630296.html. Accessed 26 Feb 2021

Tabatabei F, Wells D, in *Open Source Intelligence Investigation - From Strategy to Implementation*, ed. by B. Akhgar, P.S. Bayerl, F. Sampson (Springer, 2016), pp. 213 – 231

Tarakeswar K, Kavitha D (2011) J Comput Appl 4(1):2011

The Bellingcat Collective (2016) About Bellingcat. https://www.bellingcat.com/about/. Accessed 25 Feb 2021

The Duden. googlen. https://www.duden.de/rechtschreibung/googeln. Accessed 13 Mar 2021

The Duden. Wie kommt ein Wort in den Duden? https://www.duden.de/ueber_duden/wie-kommt-ein-wort-in-den-duden. Accessed 13. Mar 2021

The European Parliament and Council of the European Union (2016) Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation). https://eur-lex.europa.eu/eli/reg/2016/679/oj. Accessed 20 Mar 2021

The Federal Constitutional Court. The Court's Duties. https://www.bundesverfassungsgericht.de/EN/Das-Gericht/das-gericht_node.html. Accessed 14 Mar 2021

The Wayback machine. https://web.archive.org/. Accessed 21 Mar 2021

Thompson M (2016) DarkNet terminology: Definitions of the DarkNet, the Dark Web, and the Deep Web. https://coar.risc.anl.gov/coar-attends-department-of-homeland-security-hosted-darknet-summit/. Accessed 25 Feb 2021

Tweetbeaver. https://tweetbeaver.com. Accessed 21 Mar 2021

Twitter biography changes. http://spoonbill.io. Accessed 21 Mar 2021

True People Search. truepeoplesearch.com. Accessed 21 Mar 2021

W. Contributors (2021) Linkedin — Wikipedia, The Free Encyclopedia. https://de.wikipedia.org/w/index.php?title=LinkedIn&oldid=208327228. Accessed 26 Feb 2021

WhoIs Online Service. https://www.whois.com. Accessed 21 Mar 2021

Wikipedia Contributors (2021) Dark Web — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Dark_web&oldid=1011797384. Accessed 17 Mar 2021

Wikipedia Contributors (2021) Deep Web — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Deep_web&oldid=1011081276. Accessed 17 Mar 2021

Yandex. yandex.com. Accessed 21 Mar 2021