

Open Source Intelligence (OSINT) Issues and Trends

Tomislav Dokman

Faculty of Humanities and Social Sciences University of Zagreb, Croatia
dokman.tomislav@gmail.com

Tomislav Ivanjko

Faculty of Humanities and Social Sciences, University of Zagreb, Croatia
tivanjko@ffzg.hr

Summary

Open Source Intelligence (OSINT) is an intelligence product which has been processed, analysed and obtained from the publicly available information. It should be actionable and disseminated in a timely manner to the appropriate audience. Open source intelligence transfers specific knowledge to beneficiaries for them to use it in their actions and the decision-making process. Even though intelligence has for years been considered a state activity, a new postmodern intelligence paradigm (resulting from the post-Cold War period), along with the new security threats and technological improvement within the information and communication technology changed the understanding of intelligence and open source intelligence. It is believed that the bulk of intelligence, according to some estimates as much as 80 percent, comes from publicly available sources, which unquestionably indicates that this specific intelligence knowledge can be created outside of the traditional intelligence environment. The positive characteristics of publicly available information is easy availability and velocity, variety of topics, the ethical component and the low cost of its collecting; while on the other hand, in the negative context, it is characterized by manipulative character, disinformation potential, fragmented truth, contradiction and mistrust. Since public domain is a suitable platform for spreading influence, various actors in this domain are trying to have an impact on the information, making it difficult to evaluate, compare and analyse. This paper tries to explore the key advantages and disadvantages of open source intelligence in the context of intelligence activities.

Key words: open source intelligence, information, intelligence, knowledge, disinformation

Introduction

Open source intelligence (OSINT) is a gathering intelligence discipline which consists of collecting raw data accessible to everyone from publicly available sources. By further processing and analysing, this raw data is turned into actionable intelligence knowledge; or in other words, into an intelligence product, which is a crucial segment of intelligence activity. The end intelligence product, most often called intelligence information, is at the same time the crown of complex quality intelligence work since, by timely classification and distribution of intelligence knowledge to its ultimate users, the decision-making process is simplified, made quicker and easier, not only when it comes to various political decisions but also to a variety of business decisions. The traditional intelligence process is finished with this act. What preceded it was getting a request, fact finding, processing and analysis, followed by the beginning of a new cycle. It is also interesting to note here that up to date, there is no concrete record of the first use of open source intelligence. It has been claimed that the term itself, as well as the activity of collecting intelligence from publicly available sources, have by now probably been in use for hundreds of years (Hassan, 2018). Despite the fact that for a long time intelligence activity was considered to be an exclusive function of the state, the new postmodern views that came into being as a consequence of the period after the Cold War, new security challenges and technological progress within the framework of the information and technology channel, changed the understanding of the concept of both intelligence and open source intelligence. The fact that in the course of the World War II, open source intelligence was predominantly associated with the state

bears witness to that. Today, the creation of actionable knowledge using open source intelligence tools constitutes an essential quality of various non-state organizations and business sector corporations. Namely, it is against the law and it is a punishable activity for the private sector to collect strictly confidential and/or classified information by means of industrial espionage (Hulnick, 2002: 566). Hence, the creation of new knowledge by using open source intelligence activities has recently been conducted exclusively by non-state participants. It is considered that most intelligence knowledge, according to some estimates even up to 80%, comes from public sources (Wallner, 1993; Hulnick, 2002; Riley et al., 2005), which is undeniable proof that actionable knowledge can be created from publicly available information. In other words, we can discern the level of reach of open source intelligence and its role in the decision-making process. Although it is the non-state actors that have predominantly used the public knowledge, it is wider professional and academic communities that have a controversial view of the part and nature of opensource intelligence in the intelligence work of state agencies while creating an intelligence product by means of information obtained from public sources. A part of the academic and professional community considers the gathering of secret information the quality of intelligence work and intelligence knowledge only (Random, 1958; Warne, 2002; Johnston, 2005; Gil, Phythian, 2012, Warner, 2009). It needs to be pointed out that the authors listed here are primarily experts, who consider the creation of intelligence knowledge from their own specific paradigm, unfoundedly depreciating open source intelligence and giving support to the superiority of classified sources. It is Liaropulus (2006: 8) who points out that in the recent past, information was not sufficiently accessible and it was mainly a product of covert and expensive operations, while in the contemporary world, owing to the Internet, all this information is easily available, cheap and more tangible. Because of this, in this paper I would like to explain what open source intelligence is, what the open source intelligence cycle is, as well as explore key advantages and disadvantages of open source information.

Open source intelligence and the intelligence cycle

When dealing with intelligence, the term is above all determined by a definition controversy, since in its wider sense, it is identified through organization, knowledge, information and process; while in its narrower sense, the idea of intelligence refers to analysed information as the basis for decision-making. On the other hand, Tudman (2002) states that the term intelligence refers to the knowledge one has about the enemy, the activity of gathering and the organization that deals with intelligence activities, including the procedure of collecting, processing, analysing, connecting and evaluating the information at our disposal. In other words, there is intelligence potential in all collecting intelligence disciplines, no matter what the origin of data collected is. It is crucial that the data collected is structured, processed, evaluated, correlated, analysed and timely disseminated to their end users. The actionable character of an intelligence product or information is among its essential qualities.

The development of the Internet has, in our contemporary world, together with the availability of information in the digital form, had considerable impact on the growth in the quantities of open source information. The way information is gathered has also changed – until recently, it was available mostly in the printed or audio-visual form. On the other hand, with technological progress and the appearance of the Internet, an opportunity was created to gather information desired at a greater speed, no matter whether the content was displayed in one's mother tongue or a foreign language, from another country or another continent. According to Mercado (2004), it is with one movement and one click of the mouse that "analysts and officials understand the world." In other words, knowledge has never before been within such an easy reach, it has never been as widespread as today. As for defining the term, one of the founders and one of the people who have done the most on advancing open source intelligence information, Robert David Steele, emphasises that OSINT is "unclassified information that has been deliberately discovered, discriminated, distilled and disseminated to a select audience in order to address a specific question (Steele, 2006: 129). On the other hand, Williams and Blum (2018: 8) give a definition of OSINT as "publicly available information that has been discovered, determined to be of intelligence value, and disseminated by a member of the IC." "Open source intelligence is also a specific intelligence collection discipline such as human intelligence (HUMINT), intelligence gathered by interception of signals (SIGINT), imaginary intelligence (IMINT) and the scientific and technical processing of data collected from

different moving and immovable sources (MASINT) or measurement and signature intelligence (NATO 2001).” In other words, it is a discipline whose predominant feature is the availability of unclassified information. According to Mercado (2005) “open sources often equal or surpass classified information in monitoring and analysing such pressing problems such as terrorism, proliferation, and counterintelligence.” For Mercado (2005), the space of public and secret information is often interwoven with frequent meandering from the clandestine space to the public area and the other way around. The author states that it is in the public space that we often find information obtained by leaking secret information, while state systems are characterised by obtaining intelligence information from secret sources and are quite often an amalgam of open source information. It is essential to point out here that OSINT “products can reduce the demands on classified intelligence collection resources.” (Steele, 2006: 129) Steele is of the opinion that collecting data in a secret way should be reduced to a minimum, or that they should be gathered in such ways only when we know what it is that is still missing in a certain puzzle. Steele in fact here deals with the problem of optimizing state resources, with quality planning being a precondition for that, which in a concrete case means collecting the publicly available, and only after that the less readily accessible, classified information. Matey (2005: 8) thinks that “OSINT is changing the traditional conception of intelligence”, and as a consequence, there will be more and more situations where the private sector enters the field of intelligence and the process of creating intelligence knowledge. Today, there are numerous organizations which, by using open sources advance intelligence knowledge and create analytical work in the area of national security based on publicly available information. Some of the better-known ones are the American think-tank organization, RAND Corporation, along with Jane’s Information Group, and the British BBC Monitoring. The traditional open source intelligence process (request - collection - processing - analysis - dissemination) shows the cyclically interconnected parts exchanging within a circle. It is a never-ending process; after the dissemination of intelligence knowledge in the form of information is done, it goes on and a new cycle of collecting unclassified data begins. The process is initiated with a previously defined request, or an ad-hoc task, and it is with this aim that raw data is collected again. When consulting open sources, raw data is gathered from the Internet, i. e. various web pages, blogs, social media, social networks, and by accessing the traditional media such as television, radio and newspapers, including grey literature and geographical data (Hassan, 2018). Before open source information is disseminated, it is of key importance to structure, evaluate, explain and correlate the data collected so far, which is done in the data-processing phase, in the phase of analysis, when the end intelligence product is made, prior to its dissemination to the end users. In order to fully comprehend the value of OSINT, it is of crucial importance to point out the American view of it, according to which “information does not have to be secret to be valuable” (CIA, 2010), in other words, an open source piece of information can be equally useful, assuming it is obtained in time and is actionable in its character. In such cases, open source data is collected, mostly raw, the data is then sorted out, structured or processed, thus making them adequate for analytical processing and the creation of an intelligence product for dissemination to ultimate users.

The advantages and disadvantages of OSINT

In order to discuss the positive and negative features of open source information, we need to point out that the value and usefulness of this particular collection discipline outweighs its certain disadvantages and the negative traits of this discipline. These features of OSINT are summarized in Table 1 and Table 2.

Table 1. The advantages of OSINT

Advantage	Characteristics
The simplicity of obtaining information	Technological development and the Internet have resulted in simple ways of obtaining open source information. To get open source information, one needs a PC and Internet access. “This is the reason why OSINT is more accessible, ubiquitous, and valuable.” (Mercado, 2004: 47)
The speed at which a great quantity of information is gathered	Digital data are suitable for generating great quantities of information. They also make it possible to search through a great volume of available

	open source information within a very short time-span, which is not a feature of any other intelligence discipline.
Covering a wide spectrum of content units	The advantage of open source work is the coverage of a variety of national security topics, which is not possible in covert information gathering by using one human source or other intelligence collection disciplines.
Searching for data from the recent past	This discipline is superior when it comes to the investigation of post festum events, when it is necessary to collect a wider information pool about an event from the past, and investigate and/or a reconstruct a certain, especially recent, past event.
Low level of danger (safe model)	Here we are dealing with a risk-free way of information gathering, which is why this discipline can be described as the most acceptable data-collecting technique, as far as matters of security are concerned. In other words, exploiting open source information access can diminish or optimize the necessity to use much more dangerous and much more complex ways to gain factual knowledge with the help of human sources.
Open- source pieces of information do not cost much	Considering the fact that open source data is publicly available, mostly they do not have to be paid for. Still, there are certain magazines, studies and articles in think-tank organizations that can be used on condition they are given an extra fee.
Real time information gathering	This particular data gathering discipline makes it possible for us to keep track of certain events and phenomena in real time. It is possible to follow migrants' routes, investigate the consequences of a terrorist attack, outburst of an epidemic, the organization of violent demonstrations and the like.
Availability of information from various languages	Another advantage of this discipline is also obvious when it comes to looking through content in various foreign languages and having the chosen content translated into our mother tongue by means of machine translation.
Collecting information from a single spot	This secret data collecting discipline from human sources is characterised by the fact that information can be collected without going to some place, environment or society, i. e. the required data is collected directly from our offices.
Simple data dissemination procedures	Considering the fact that here we are dealing with the data readily available to everybody, their dissemination should not be based on spreading them restrictively or in a limited way; such pieces of information are suitable for both horizontal sharing and they can be made available to the broader public in general.
Secret information authentication/ verification	The comparison of classified information with publicly available data is an outstanding method of checking whether there is maybe some "secret" data being circulated in open sources. There is a possibility that some data were previously publicly known, so this method can be used to ascertain the validity and evaluate both the content and the source of covert information, at the same time channelling further collection of classified information.
Business intelligence suitability	Using open source data can definitely be helpful in bringing adequate business decisions, planning future projects, conducting pilot projects, market research, comparing and evaluating competitors, as well as ensuring appropriate publicity for one's own business ideas.

Other authors have also been studying the good and bad sides of open source intelligence information and their opinions concur with the insights given here. For instance, Annie Ahuja (2018: 470) states the advantages of open source intelligence are: 1) Less Expensive; 2) Accessing Information; 3) Security; 4) Business; 5) Social Media; 6) Updated Data and Metadata; 7) Semantic Understanding; 8) Applications. Simultaneously, Mercado (2005) emphasises OSINT's "value advert, speed, quantity, quality, clarity, ease of use and cost." OSINT is a unique discipline that can share its product with

everyone, the profit being greater if we take into consideration the fact that it is also shared with humanitarian aid missions and forces of security, law and order. The fact that the results of OSINT's inquests are used as contributions to early warning systems also needs to be born in mind (Steele, 2006: 133). On the other hand, it is important to point out that OSINT has not been put in place to replace other intelligence disciplines in the production of actionable knowledge, i.e. it is not aimed at one discipline replacing an existing one.

Table 2. The disadvantages of OSINT

Disadvantage	Characteristics
Great quantities of accessible information	Public space is characterised by great quantities of available information. Quite often the data is badly laid out and therefore confusing, and then it is also difficult to find the exact data we are looking for. Hassan (2018) thinks that we live in an information world, with millions of people in constant communication and constantly sharing all kinds of information. Namely, the daily data production, according to some estimates, amounts to about 2.5 quintillion bytes, the assumption being that this number is bound to increase even further with further Internet growth (Marr, 2018)
Information contradiction	The availability of a wide information fund has been caused by the democratization of the net, coupled with the opportunity of an individual to create and spread information without too much difficulty. It is important to objectively view the data collected in order for the product of public intelligence knowledge to be as accurate as possible.
The manipulative character of publicly available information	Public space is obviously an excellent area for spreading disinformation and fake news in order to achieve certain effects. By influencing the corpus of public knowledge, using information and counter information in the public arena, those in charge of various campaigns want to gain information superiority (Akrap, 2009: 80). It is due to this reason that this discipline is characterised by the constant checking and rechecking of all the material at our disposal.

Conclusion

Until the emergence of the Internet and the state-of-the-art communication and information-sharing applications, open source intelligence information was considered to be of low value as an intelligence discipline. Not accepting the real value of open source intelligence mostly corresponded with an a priori attitude that an intelligence product could exclusively come from covert information sources, while work with publicly available data was considered to be less valuable and less interesting for the intelligence activity. However, in the current, above all changing security environment, when traditional threats became more divergent and changed in their configuration, especially when the gathering of secret data became a long-lasting, expensive and complex process, intelligence services began creating open source intelligence knowledge. Besides that, this collecting discipline also proved to be an invaluable source of knowledge apart from state activity and has, therefore, been fully assimilated into other spheres of public life. Moreover, the intelligence activity, including open source intelligence information, is no longer only a feature belonging to the state itself. Gathering, analytical processing and the production of relevant and actionable intelligence knowledge are now typical for both the state and non-state actors. However, it is envisaged that in the future, OSINT will be faced with great challenges caused by immense quantities of unstructured data, because in order to collect them and sort them out appropriately, an ever-increasing development of sophisticated software will be required. Since we are here dealing with the greatest source of information of all intelligence disciplines, great effort will be necessary to extract quality and timely information from a considerable number of sources, in order to attain usable actionable knowledge. At the same time, it needs to be pointed out that public space is an area for spreading influence, propaganda and a variety of interests, which is why the checking of information has to be the main feature of work when dealing with the publicly available content.

References

- Akrap, G. (2009). Informacijske strategije i oblikovanje javnoga znanja. // *National Security and the Future* 10, 2, 77-151
- Annie, A. (2018). A Review Paper on Open Source Intelligence: An intelligence sustenance. // *International Journal of Recent Trends in Engineering & Research* 4, 4, 463-474
- Best, C. (2007). Open Source Intelligence. // *Mining Massive Data Sets for Security: Advances in Data Mining, Search, Social Networks and Text Mining, and their Applications to Security* / Fogelman-Soulie, F., Perrotta, D., Piskorski, J., Steinberger, R. (eds.). Amsterdam: IOS Press, 331-344
- CIA. (2010). INTelligence: Open Source Intelligence. 2010/07/23. <https://www.cia.gov/news-information/featured-story-archive/2010-featured-story-archive/open-source-intelligence.html> (2019/08/05)
- Hassan, N. (2018). An Introduction To Open Source Intelligence (OSINT) Gathering. 2018/08/12. <https://www.secjuice.com/introduction-to-open-source-intelligence-osint/> (2019/08/10)
- Hulnick, A. S. (2002). The Downside of Open Source Intelligence. // *International Journal of Intelligence and CounterIntelligence* 15, 4, 565-579
- Johnston, R. (2005). *Analytic Culture in the US Intelligence Community: An Ethnographic Study*. Washington: Central Intelligence Agency
- Liaropoulos, A. N. (2006). A (r)evolution in intelligence affairs? In search of a new paradigm. Athens: Research Institute for Europe and American studies (RIEAS)
- Marr, B. (2018). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. 2018/05/21. <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#12e73d5f60ba> (2019/08/11)
- Mercado, S. C. (2004). Sailing the Sea of OSINT in the Information Age: A Venerable Source in a New Era. // *Studies in Intelligence* 48, 3. <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol48no3/article05.html> (2019/08/15)
- Mercado, S. C. (2005). Reexamining the Distinction Between Open Information and Secrets. // *Studies in Intelligence*. 49, 2. https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/Vol49no2/reexamining_the_distinction_3.htm (2019/08/15)
- Matey, G. D. (2005). Intelligence Studies at the Dawn of the 21st Century: New Possibilities and Resources for a Recent Topic in International Relations: UNISCI Discussion Papers. May. <https://www.ucm.es/data/cont/media/www/pag-72533/Gustavo2.pdf> (2019/08/11)
- NATO (2001). *Open Source Intelligence Handbook*. November 2001. <https://archive.org/details/NATOOSINTHandbookV1.2> (2019/08/19)
- Random, R. A. (1958). Intelligence as a science. 1958, last updated 2011. https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/vol2no2/html/v02i2a09p_0001.htm (2019/08/10)
- Riley, J. K., Treverton, G. F., Wilson, J. M., Davis, L. M. (2005). *State and Local Intelligence in the War on Terrorism*. Santa Monica: RAND Corporation
- Steele, R. D. (2006). Open Source Intelligence. // *Handbook of Intelligence Studies* / Johnson, Loch K. (ed.). New York : Routledge, 129-147
- Tudman, M. (2002). Informacijska znanost i izvjesnice. // *Informatologia* 35, 4, 244-251
- Wallner, P. F. (1993). Open Sources and the Intelligence Community: Myths and Realities. // *American Intelligence Journal* 14, Spring/summer, 19-24.
- Warner, M. (2002). Wanted: A Definition of Intelligence: Understanding Our Craft. // *Studies in Intelligence* 46, 3. <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol46no3/article02.html> (2019/08/11)
- Warner, M. (2009). Intelligence as risk shifting. // *Intelligence Theory: Key questions and debates* / Gill, P., Marrin, S., Phythian, M. (eds.). London: Routledge, 16-32
- Williams, H. J., Blum, I. (2018). *Defining Second Generation Open Source Intelligence (OSINT) for the Defense Enterprise*. Santa Monica: RAND Corporation