Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
May 2, 2019
A Regional Conference of the Society for Industrial and Systems Engineering

Closing the Gap between Data and Open Source Intelligence

Jacob Hedges, Sam Messina, Danielle Peck, Sierra Rosdahl, and Steven Song

Department of Systems Engineering and Mathematical Sciences,
United States Military Academy, West Point, NY

Corresponding author: jacob.hedges@westpoint.edu

Author Note: Cadets Jacob Hedges, Sam Messina, Danielle Peck, and Sierra Rosdahl are seniors in the Department of Systems Engineering at the United States Military Academy. Major Steven Song, Assistant Professor in the Department of Systems Engineering, is the capstone advisor and provided guidance on the research project.

Abstract: Due to the massive influx of technology and data, the demand for data exploitation and analysts have increased. The Army Open Source Office (AOO) provides training and tools in order to provide operational capabilities for Army intelligence analysts. However, the United States Army's policies and regulations play a significant role in data exploitation and collection. This paper addresses the challenges, limitations, and identifies the gaps within open source intelligence (OSINT). The research highlights a potential solution and recommendation using a systems approach. First, a value model which determines the overall value of current OSINT tools. Second, it is recommended that a platform is created which will house both future and current technologies. This will ultimately streamline the AOO's strategies and drive future practices. The research project provides the AOO the opportunity to increase effectiveness, decrease unnecessary training, and reduce costs. As a result, the proposed value model supplements the AOO's ability to evaluate potential tools and helps close the information gap by providing a detailed assessment of the AOO's current toolset.

Keywords: Open-Source Intelligence, Value Model, Publicly Available Information

1. Introduction

The Army OSINT Office lacks a value model which evaluates current OSINT tools against overarching requirements to identify gaps, assess against emerging requirements, and to minimize overlap in capabilities. The research identifies the AOO lacks a quantifiable method to evaluate their current tools and assess future technologies while aiming to address the AOO's requirements. Thus, leading to three research questions for further analysis:

(1) What tools does the AOO currently have that the AOO does not need?
(2) What tools does the AOO not currently have, but would be helpful in the future?
(3) What future technology would be beneficial to the Army's OSINT capabilities?

Since the AOO focuses on providing capabilities for the analyst, the goal is to create a value model in order to determine the tools that provide the most value to the organization.

The systems approach is formally known as the Systems Decision Process (SDP) (Parnell et al., 2011, p.16). The research determined that the AOO needs a way to systematically assess current OSINT tools. A model using a Likert Scale to determine total value scores of each system was implemented to address these concerns. This model optimizes the process of evaluating current and potential tools. With better defined system requirements, the AOO can provide appropriate tools and exploit data more efficiently.

2. Literature Review

According to "The Road to the Data Strategy for Army Intelligence," data is processed in four steps: plan and direct, collect and process, produce, and disseminate (Brustman, Christensen, Russo, Edmiston, & Saddler, 2018). These steps are utilized by the Army intelligence community. Furthermore, big data consists of five characteristics: volume, variety, velocity, veracity, and variability (Brustman, Christensen, Russo, Edmiston, & Saddler, 2018). Since there are not enough analysts trained to review all the data, Army intelligence analysts are unable to process and analyze the volume of data at the rate it is collected. Therefore, the big data challenge presents a threat to the United States' national security due to the significant amount of unprocessed data.

The Army, in particular, has a vested interest in big data, especially with countries that are perceived threats to the United States. For example, the Islamic State of Iraq and al-Sham (ISIS) uses social media platforms to distribute powerful, emotional images around the world (Farwell, 2014). This extremist group targets individuals on social media networks, in an attempt "to recruit fighters and intimidate enemies" (Farwell, 2014).

According to *The Intelligencer*, open source intelligence is "the collection, processing, analysis, production, classification, and dissemination of information derived from sources and by means openly available to and legally accessible and employable by the public in response to official national security requirements" (Schaurer, Störger, 2012). Many government agencies such as the FBI, CIA, and the Army's Intelligence and Security Command (INSCOM) conduct OSINT research to enhance intelligence beyond the foundational level. One of INSCOM's mission is to train analysts to provide a strategic and tactical advantage to friendly forces using reliable and open source research through the AOO (INSCOM, 2018). Additionally, the AOO provides policy guidance to Army OSINT analysts, OSINT tools, and develops OSINT capabilities.

OSINT provides relevant data to intelligence communities, both in and outside of the Army. Although multiple intelligence agencies use OSINT because it is cost effective, analysts must pay for data access and specific analytic tools in order to accomplish their mission (Paulson, 2008). In comparison to other intelligence sources, OSINT analysts are able to access publicly available information (PAI) at any time, making the data more readily available and easier to access (Paulson, 2008). The Army relies on OSINT for its timely, cost effective information. For example, Digital Forensic Intelligence uses OSINT because it is "fast, flexible, dynamic, communicable, shareable, [and] partner forming" (Quick, Choo, 2018). While the Army OSINT community strives to quickly exploit PAI, they must abide by specific policies. Specifically, the Army must abide by any other legal policies that pertain to intelligence including the Privacy Act of 1974, as stated in DoD Manual 5240.01 (Department of Defense Manual 5240.01, 2016). Soldiers that conduct OSINT research must be aware of the policies that govern intelligence in order to prevent repercussions to the individual, the agency, and even the mission.

Limitations that apply to OSINT include the systems' inaccuracy and policies related to data collection. Therefore, OSINT is not always valid or reliable. One of the jobs of an analyst is to sort the incoming information and determine the validity of the resource. Additionally, OSINT works directly with other sources of intelligence, such as counterintelligence (CI), human intelligence (HUMINT), and signal intelligence (SIGINT), which limits OSINT to only PAI (ATP 2-22.9, 2017). AOO's limitations are defined by Army doctrine ATP 2-22.9 (ATP 2-22.9, 2017). This ATP states that once data is collected, it is no longer considered OSINT (ATP 2-22.9, 2017). Collection, "includes information obtained or acquired by any means, including information that is volunteered to the component" (ATP 2-22.9, 2017). According to this doctrine, OSINT analysts are only allowed to view PAI, but cannot copy, save, or supplement information (ATP 2-22.9, 2017). The language barrier is also an important factor when filtering through big data (Quick, Choo, 2018). This means that data must be translated before it can be exploited and filtered. Having computer systems that can accurately translate big data increases the rate at which analysts can provide intelligence.

One of the challenges with OSINT stems from social media platforms. AOO usually uses commercial off-the-shelf tools (COTS) for gathering PAI and big data (Vaughn, 2015). This section focuses on the current tools OSINT is utilizing: S-1, D-1, D-2, RF-2, Z-1, B-1, O-1, F-1, S-2, CH-1, and M-1 (Army Open-Source Office, 2018). Most of these tools listed above were developed for commercial purposes: advertising on social media platforms, commercial brand management, and data gathering on users. However, the current tools used do not satisfy all of the OSINT needs because most of them are made for commercial use. Despite this imperfect fit, the AOO can still use these current tools to help achieve their mission.

3. Methodology – The Systems Decision Process (SDP)

SDP is a continuous, four-stage process that is conducted to produce a refined solution that meets the stakeholder's needs. The four phases consist of problem definition, solution design, decision making, and solution implementation with systems thinking as the underlying foundation. The SDP aided the systems approach to understand and analyze AOO's challenge with OSINT. The SDP is "a collaborative, interactive and value-based decision process that can be applied in any system life cycle stage" (Parnell et al., 2011).

3.1 Problem Definition

First, the problem definition phase establishes the scope of the issue. This phase contains research and stakeholder analysis, functional requirements analysis, and value modeling (Parnell et al., 2011). The goal of this stage is to have a clearly defined problem statement that aligns with the stakeholders' needs. The initial problem statement for AOO stated: Conduct a holistic analysis of the existing technologies and tools for exploiting PAI and provide a framework and methodology for evaluating existing and emerging technologies.

After conducting the literature review, the stakeholder analysis helped identify the objectives, functions, and constraints of a system. The research applied systems thinking to understand the environmental factors affecting a system and identifying the relevant stakeholders. To better understand the problem, the research obtained inputs from stakeholders and subject matter experts (SME) using interviews, focus groups, and surveys. The main stakeholders of the project included the Director, Deputy Director, and the Senior OSINT instructor of the AOO. Despite the limited size of the AOO, the research team collected seven survey results. These results were crucial for defining the seven attributes of the model and determining their correlating swing weights.

Our research identified seven attributes to assess the effectiveness and importance of current and future AOO tools. These seven attributes included security, cost, interoperability, speed, ease of use, adaptability, and accuracy. Security includes measures taken to mask identity (e.g. password protection, secure browsing). Cost is the amount paid to procure OSINT tools (e.g. Cost of program, training, and maintenance). Interoperability is defined as the ability to operate with other systems (e.g. interchangeable parts, ability to connect and communicate with other platforms, ability to share and aggregate information). Speed is the rate of performance (e.g. time in which it takes to process and disseminate information). Ease of use is described as the ability of the analyst to operate the tool with minimal guidance or training (e.g. accessibility, intuition). Adaptability is the ability to make a tool fit a different function or be modified for a specific function (e.g. the ability of the system to gather data from multiple sources and change with the dynamic aspects of the internet). Finally, accuracy is the freedom from mistake or error (e.g. ability to filter our false information). These definitions provided a better understanding for assessing and evaluating OSINT tools.

The survey requested the stakeholders to rank the seven attributes based on their understanding of the tools. After the surveys, five stakeholders were interviewed. The interviews focused on the three research questions and better define the capabilities of each tool. The findings from our surveys demonstrated that security, interoperability, ease of use, and adaptability were the key attributes the AOO valued most.

Functional analysis provides an understanding of the system's relationships in a hierarchical form. The functional hierarchy provides a clear understanding of the functions the system will perform and serves as the foundation for the candidate solution design. The functions and subfunctions, within Figure 1, depict the seven key attributes for evaluating OSINT tools. After completing research, stakeholder analysis, surveys, and interviews, the revised problem statement is defined as: The AOO lacks a mechanism to evaluate current toolsets to identify gaps, assess against emerging requirements, and to minimize redundant capabilities.

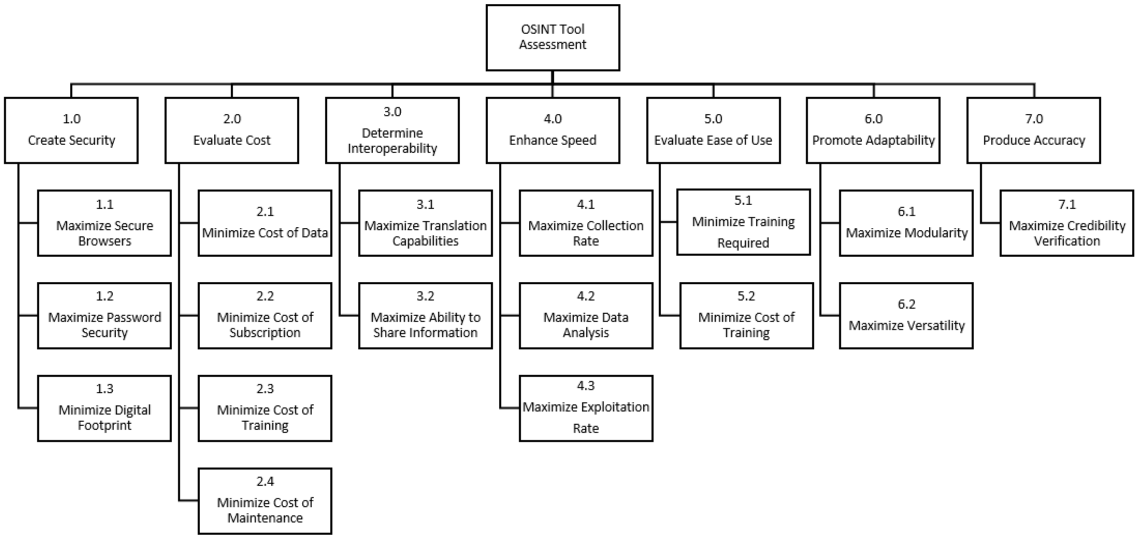
OSINT Tool Assessment
1.0 Create Security
1.1 Maximize Secure Browsers
1.2 Maximize Password Security
1.3 Minimize Digital Footprint
2.0 Evaluate Cost
2.1 Minimize Cost of Data
2.2 Minimize Cost of Subscription
2.3 Minimize Cost of Training
2.4 Minimize Cost of Maintenance
3.0 Determine Interoperability
3.1 Maximize Translation Capabilities
3.2 Maximize Ability to Share Information
4.0 Enhance Speed
4.1 Maximize Collection Rate
4.2 Maximize Data Analysis
4.3 Maximize Exploitation Rate
5.0 Evaluate Ease of Use
5.1 Minimize Training Required
5.2 Minimize Cost of Training
6.0 Promote Adaptability
6.1 Maximize Modularity
6.2 Maximize Versatility
7.0 Produce Accuracy
7.1 Maximize Credibility Verification

Figure 1. Functional Hierarchy of OSINT Tool Assessment.

3.2 Model Development

Creating a model will ultimately provide the AOO with a systematic process to evaluate their tools. After constructing the functional hierarchy, a qualitative model was created. The majority of the value ratings used to judge current and prospective tools were qualitative. These value ratings were created based on stakeholder survey feedback. The research team used this feedback to assign swing weights to the attributes in the working model. The higher the swing weight the greater the impact on the model's output for determining the effectiveness of the tool. The opposite is true for a lower swing weight. In order to convert qualitative survey information to quantitative values, an assessment rubric was created using a Likert Scale. The assessment rubric is flexible by design and will allow AOO to use the constructed scale for rating tools of interest.

The overall value for the OSINT analysis tool was found by summing the value of the individual attributes of the tool. The values of the tool attributes were calculated by using the numeric output from the Likert Scale as an input for the specific qualitative function and multiplying the output by its scalar swing weight. The equation below depicts this process.

$$V_t = \sum s_q f_{qt}(n) \; where \; V_t = value \; of \; the \; tool; \; s_q = swing \; weight; \; f_{qt}(n) = rating \; of \; the \; tool \tag{1}$$

The values from this equation were inputted into a value matrix to obtain the overall value of each tool. These results are displayed in Figure 2. Based on the results, the analysis determined the sensitivity of each swing weights' value. The raw value scores for each tool's attributes were derived from a non-linear value curve. Each value curve is unique to its characteristic and was developed based on the stakeholder's needs and requirements. Overall, the value model served as a platform, which converts non-numerical data into objectively comparable total value scores.

4. Results & Analysis

After evaluating the current OSINT tools, 33 emerging platforms were identified as potential candidates for the AOO's inventory using the systems requirements shown in Table 1. Each future tool was scored using the Likert Scale rubric within the model to see if the AOO could utilize these technologies. The model can objectively assess how the potential tools compare to the current tools. Research revealed that there were technologies available that could benefit the AOO. The systems requirements were developed using the valuable insights were gained through research and assessing new technologies. However, some platforms of the tools were relatively undeveloped. Undeveloped tools could be adapted for the AOO using the systems requirements for emerging technologies. Our research suggests that AOO could optimize their OSINT tool portfolio using the system requirements when evaluating future technologies.

Table 1. System Requirements for Emerging Technologies

United States Based	The system shall utilize prospective tools based out of the United States.
Interoperable	The system shall be interoperable and seamlessly communicate with other OSINT tools.
No Required Login	The system shall not require a hard login that requires full profile account.
Transparency of Information	The system shall be able to share and disseminate vital information.
One Platform	The system shall be on one structured platform that contains multiple OSINT tools.
Dark Web Access	The system shall have dark web capabilities.
Image/Language Translations	The system shall detect and translate images, videos, slang, dialects, and traditional languages.
Security	The system shall have tools that must reach basic security requirements for government network and use.

The value model determined how 11 of the AOO tools compare to each other based on their individual value scores, as shown on the left in Figure 2. Furthermore, seven potential tools were proposed to mitigate the OSINT gaps. These seven tools were chosen out of the 33 based on specific criteria. This elimination included tools that were not U.S. owned and developed and lacked the potential to close gaps within AOO. The potential future tools consist of SocialMention (SM), World Lens (WL), Open Source Indicators (OSI), Tableau (TB), and UiPath (UP) can enhance their current operational capabilities. The figure below displays the value scores of AOO's currently used tools, on the left, and displays the seven candidate future tools, on the right. The findings in this paper are based solely on the value of the tools because the costs of AOO's tools are confidential. Therefore, cost versus value analysis is not included as part of the results.

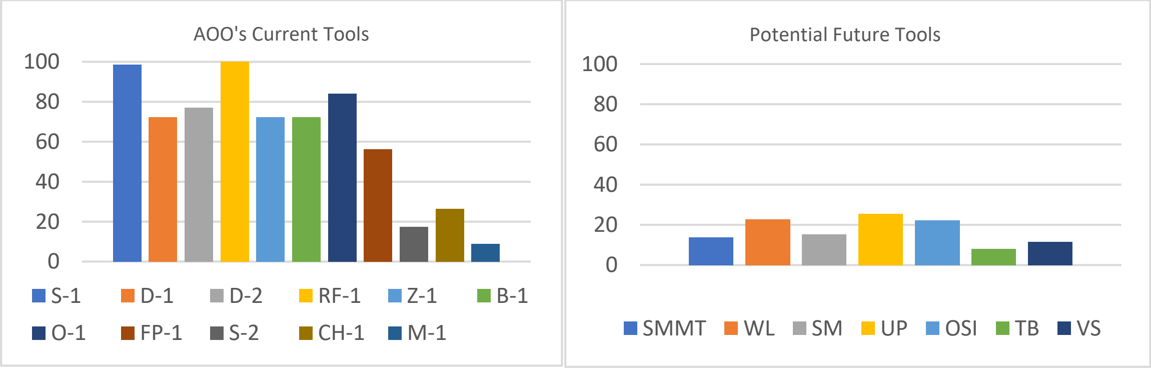

Figure 2. Value model output of the current AOO tools (left) and potential tools (right).

As shown in Figure 2 (left), three tools vastly outperformed the others. RF-1, S-1, and O-1 received high scores, which indicates these are critical for AOO's operational success. However, M-1, S-2, and CH-1 all received scores under 30. This suggests they may be non-essential for the AOO. Thus, the AOO has the opportunity to improve their toolset by replacing tools with higher value scores. Sensitivity analysis was also conducted on all seven attributes based on the current tools. As shown in Figure 3, "Ease of Use" is the most sensitive characteristic and, arguably, the most subjective. "Accuracy" is the least sensitive because it can be evaluated analytically. The sensitivity analysis provides important insight into how trade-offs and combinations can be implemented using current tools. For example, creating an open, structured platform, which contains multiple OSINT tools, not only will consolidate the tools, but also provide convenient access to all the tools. This model enables the AOO to identify gaps within their current tool portfolio, identify the most beneficial tools, and determine tools/technologies that can replace low scoring tools.

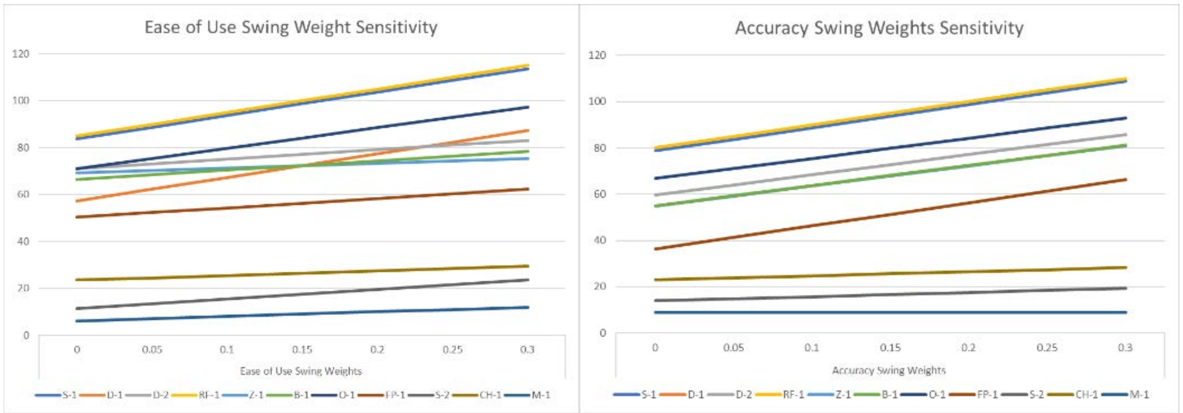

Figure 3. Sensitivity Analysis of Ease of Use (left) and Accuracy (right) for each current tool

5. Conclusion and Future Work

The value model identified the highest and lowest rated tools based on their value score, with RF-1 scoring a 100. We recommend the AOO eliminate M-1, S-2, and CH-1 from their inventory due to their low scores. Additionally, independent research was conducted to present future tools and technologies to the AOO. These findings were submitted to the AOO to further guide their decision making process. Two important takeaways from these findings included adding a tool that has image and language translation capabilities. Second, creating a future platform that contains both current and future tools will benefit the AOO and increase their capabilities. Moreover, the model provides the AOO the ability to rank emerging tools against the existing tools. The value model can serve as a starting point for evaluating emerging technologies for the Army OSINT community.

With the landscape of the internet constantly evolving, there is an ongoing need for new OSINT tools. The strict regulation applied to AOO prohibits the agency from accessing a variety of tools. Additionally, all information and intelligence gathered must be verified and validated. These factors require the AOO to keep a wide variety of adaptable tools. Based on the high rankings from the value model in Figure 2, we recommend the AOO consider implementing UiPath, World Lens, and Open Source Indicators to close the gaps in analyzing big data, and image and language translations.

Research indicates several opportunities for AOO to close the gap between data and OSINT. Our model indicates that AOO could benefit from a systemic evaluation of their portfolio. The tools should be evaluated based on the seven attributes deemed most important by the stakeholders and the system requirements. The value model's flexibility allows the AOO to evaluate new and potential tools. Based on our research of future technologies, the systems requirements can provide as a guide for organically developing OSINT capabilities. As a result, we recommend the AOO implements the value model and use the systems requirements to modernize the Army OSINT portfolio.

6. References

Army Open-Source Office. (2018). Current Tools Provisioned by the AOO in OS302.

ATP 2-22.9. Open-Source Intelligence. 30 June 2017.

Brustman, K., Christensen, E., Russo, H., Edmiston, R., & Saddler, R. (2018, July). The Road to the Data Strategy for Army Intelligence. Retrieved from https://www.ikn.army.mil/apps/MIPBW/.

Department of Defense Manual 5240.01. (2016, August). "Procedures Governing the Conduct of DoD Intelligence Activities."
Washington, D.C. https://dodsioo.defense.gov/Portals/46/DoDM%20%205240.01.pdf?ver=2016-08-11-184834-887.

Farwell, J. P. (2014, November). The Media Strategy of ISIS: Survival. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/00396338.2014.985436.

Parnell, G., Driscoll, P., & Henderson, D. (Eds.). (2011). *Decision Making in Systems Engineering and Management*. Hoboken, NJ: John Wile & Sons, Inc.

Paulson, T. M. (2008). *Intelligence Issues and Developments*. New York: Nova Science.

Quick, Darren., Choo, Kim-Kwang. (2018, January). "Digital Forensic Intelligence: Data subsets and Open Source Intelligence (DFINT+OSINT): A timely and cohesive mix". *ScienceDirect.* Retrieved from https://www.sciencedirect.com/science/article/pii/S0167739X16308639.

Schaurer, Florian, and Jan Störger. "The Evolution of Open Source Intelligence (OSINT)." *The Intelligencer*, Association of Former Intelligence Officers, 2013, www.afio.com/publications/Schauer_Storger_Evo_of_OSINT_WINTERSPRING2013.pdf.

United States Army Intelligence and Security Command. (2018). MI Professional Bulletin: Enabling Readiness for the Intelligence Enterprise (PB 34-18-3) Retrieved from https://www.ikn.army.mil/apps/MIPBW/MIPB_Issues/MIPBJul_Sept18FinalIKN2.pdf

Vaughn, A. (2015, February 09). Mil-COTS Power Supplies, COTS Power Supply. Retrieved from http://aegispower.com/index.php/blog/180-what-are-mil-cots-power-supplies.