Addressing Challenges in a Dangerous World: Developing a Design Science Artifact for Advancing Open Source Intelligence (OSINT) Research

Franz Kayser
ESG
franz.kayser@esg.de

Dr. Thomas Mayer
ESG
thomas3.mayer@esg.de

Prof. Dr. Michael Bücker
University of Applied Sciences Münster
michael.buecker@fh-muenster.de

Abstract

Open Source Intelligence (OSINT), which derives intelligence from public data, has gained attention since the Russian invasion of Ukraine. Despite attempts at standard definitions, research on technology-driven intelligence gathering remains ambiguous. This paper employs a Design Science Research (DSR) approach to categorize this construct. Analyzing sixty studies through a structured literature review, three domains were identified: maturity, Intelligence Cycle (IC) phase, and use case. The resulting framework, developed into a trend radar (TR), was evaluated with expert interviews, revealing technological gaps in the planning/direction and dissemination/integration phases. Although intelligent support technologies exist, practical implementation lags behind theory, with the human factor remaining central to OSINT. Findings suggest future research should focus on developing applications for underserved phases and investigating why proven solutions are not widely adopted, considering legal, ethical, political, and social factors. This study contributes to the literature by providing a knowledge base, identifying research gaps, and guiding further research.

1. Introduction

OSINT, the process of extracting intelligence from public data (Dos Passos, 2017), has gained attention, especially since the 2022 Russian invasion of Ukraine. Real-time analysis of social media has proven pivotal in uncovering valuable insights (Smith-Boyle, 2022). Despite numerous attempts to define OSINT (Pastor-Galindo et al., 2020; Yogish and Krishna, 2021), controversy persists due to ongoing advancements in computer and data sciences that continuously enhance collection and analysis capabilities (Ghioni et al., 2023).

The surge in open communication channels has led to an 'information explosion' (Hwang et al., 2022), making previously restricted data publicly accessible (Williams and Blum, 2018) and reshaping intelligence paradigms (Dokman and Ivanjko, 2020). However, fundamental scientific literature in the field remains limited (Herrera-Cubides et al., 2020), failing to keep pace with rapid developments (Ghioni et al., 2023; Williams and Blum, 2018). Key questions regarding the existence of autonomous third-generation OSINT systems remain unanswered (Ghioni et al., 2023; Pastor-Galindo et al., 2020), and OSINT use cases unexplored (AlKilani and Qusef, 2021; Ghioni et al., 2023), largely due to the challenges in accessing government and security sectors (Herrera-Cubides et al., 2020; Pastor-Galindo et al., 2019).

A robust, structured framework is needed to enhance the scientific understanding and application of OSINT. This study addresses the research question: *How can current OSINT trends, in the form of technologies and their characteristics, particularly maturity levels and use cases, be presented in a unified way that bridges academic research and practical applications?* This paper investigates current OSINT trends using the DSR approach (Peffers et al., 2007). The method involves a systematic review (Cleven et al., 2009) to analyze OSINT literature, establishing a unified knowledge base. OSINT technologies and their characteristics will be visualized in a TR, evaluated through semi-structured interviews (Gläser and Laudel, 2009) with German security experts, and analyzed via qualitative content analysis (Billings, 1997). The research background material is available on GitHub (Kayser, 2024).

URI: https://hdl.handle.net/10125/109494
978-0-9981331-8-8

2. Theoretical Background

2.1. Open Source Intelligence

One of the earliest referenced definitions was published by the North Atlantic Treaty Organization (NATO (2001)): *"OSINT is information that has been deliberately discovered, discriminated, distilled, and disseminated to a selected audience, [...], in order to address a specific question. OSINT, [...] thus applies the proven process of intelligence to the broad diversity of open sources [...] and creates intelligence."* However, today the discipline is no longer seen as a purely governmental matter. Private research institutions and organizations (Böhm and Lolagar, 2021) are also massively driving the development of such systems (Dokman and Ivanjko, 2020). The focus is shifting to developing OSINT into third-generation robust, autonomous solutions (Pastor-Galindo et al., 2019).

2.2. Intelligence and Intelligence Cycle

OSINT's core task is to generate decision-making intelligence (NATO, 2001). The generation process of such an intelligence product is referred to as the IC specified by the Central lntelligence Agency (CIA, 1987). It represents the central element of every intelligence discipline (Reuser, 2017). The link between the phases is that the result of each preceding phase serves as input for the subsequent phase, following the U.S. Joint Force Command (USJFCOM, 2013), continuously iterated to meet new requirements (Gibson, 2016). Today, to represent external influences or assign responsibilities (Lowenthal, 2020; Phythian, 2013), numerous variations can be found (Reuser, 2017). The IC should hence be seen less as a guideline and more as an informal coordination element (Hwang et al., 2022). USJFCOM (2013) segmented the cycle into six phases (Figure 1).

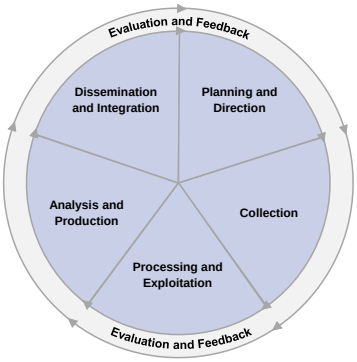

Figure 1. Intelligence Cycle (USJFCOM, 2013)

The planning/direction phase combines the identification, definition, prioritization, and monitoring of the requirements (USJFCOM, 2013). The collection phase refers to the gathering of raw data (CIA, 1987). It consists of iterative repetition of research (NATO, 2001) to make the query more precise with each run (Pastor-Galindo et al., 2020). The processing/exploitation phase involves condensing these data volumes into action-relevant information (USJFCOM, 2013). Analysis/production refers to the synthesis of the information obtained into a timely and accurate intelligence product (NATO, 2001). The final phase consists of handing over the product to the 'customer' in a usable form (CIA, 2023; Williams and Blum, 2018). Evaluation/feedback are not to be regarded as individual phases but take place continuously to achieve progressive optimization (NATO, 2001; USJFCOM, 2013).

2.3. Previous Studies

Eight publicly accessible literature reviews exist on OSINT. Dos Passos (2017) showed how big data and data science enhance decision-making. Pastor-Galindo et al. (2019, 2020) provided insights into OSINT's state, focusing on cyber security enhancements. They conducted the first rudimentary mapping of OSINT trends, observing its use in social opinion and sentiment analysis, cyber crime and organized crime, as well as cyber security and cyber defense. García Lozano et al. (2020) identified methods for computer-assisted veracity assessment of public information, while Herrera-Cubides et al. (2020) researched the production of research/educational materials. They concluded that OSINT publications are less common compared to other trending topics. Yogish and Krishna (2021) explored AI implementation in cyber security, showing its potential to simplify OSINT given increasing data volumes. Hwang et al. (2022) investigated security threats and cyber criminality through OSINT misuse. Ghioni et al. (2023) examined the political, ethical, legal, and social implications of OSINT in conjunction with AI, highlighting the absence of a comprehensive framework and the early stage of third-generation OSINT.

3. Research Methodology

The study follows the iterative DSR approach, a theory-based research paradigm for developing a directly applicable solution in the form of an innovative artifact (vom Brocke et al., 2020) to solve a (practical) problem (Peffers et al., 2007). Hence, the model is ideal for creating an artifact to address the definitional gap and lack of academic frameworks. The DSR methodology

includes six successive activities (Peffers et al., 2007, Figure 2). Section 1 summarizes step 1, while steps 2-5 will be discussed in detail in sections 3.1-3.7. Sections 4 and 5 present the outcomes of step 6. Continuous evaluation of steps 1-4 occurred throughout the study following Sonnenberg and vom Brocke (2012).

3.1. Design Objectives of the Solution

The design objectives, derived from the research question (Peffers et al., 2007), are categorized into content-related (CO) and formal objectives (FO).

CO1: Third-generation OSINT systems remain unconfirmed (Ghioni et al., 2023). The artifact must reflect the intelligence generation process to map technologies by usage to respective phases and identify research gaps.

CO2: Key use cases (AlKilani and Qusef, 2021; Dokman and Ivanjko, 2020; Ghioni et al., 2023), technologies, and their characteristics are lacking in OSINT research (Ish et al., 2022). The artifact must include technology maturity to inform research status and use cases to indicate research directions.

FO1: Given OSINT's rapid evolution (Ghioni et al., 2023), the artifact requires a simple, standardized structure to quickly identify research gaps and ensure cross-disciplinary applicability.

FO2: Due to the high field dynamic, predicting future developments is challenging (Benes, 2013). The artifact should be designed for continuous expansion to remain relevant over time.

3.2. Evaluation of the Problem Statement and Design Objectives

In section 2, the theoretical background of the research is outlined, focusing on the IC, which is crucial for the artifact's development. The artifact is evaluated for its compatibility within the OSINT framework. Additionally, the review of prior studies highlights a significant gap in comprehensive OSINT research, emphasizing the inquiry's importance and relevance.

3.3. Design and Development

The third activity is a systematic literature review, guided by Cleven et al. (2009). The taxonomy by Cooper (1988) scoped the review, setting up classification categories for concept matrices to structure the literature analysis (Webster et al., 2002). To standardize verification across the IC, general categories reflecting established technology evaluation criteria were defined for each phase, except the iterative evaluation/feedback phase. For the collection phase, six categories were developed:

- **Use case:** Application areas of the technologies.
- **Data:** Composition and types of data foundations, including data format and source.
- **Process:** Degree of automation in the technologies (manual, semi-automated, automated, fully automated; Billings, 1997; Duncheon, 2002; Endsley and Kaber, 1999).
- **Technology:** Material and immaterial means used for managing information (Bleck, 2004).
- **Technology Complexity:** Assessed through subcategories of volume, variety, and velocity (Singh and Singh, 2012).
- **Maturity Level:** According to the phases innovation, prototype, and market establishment (Stich et al., 2022).

Similar categories were adapted for the other phases of the IC, with complexity measured via the analytics spectrum: descriptive, diagnostic, predictive, and prescriptive (Delen and Demirkan, 2013). The literature search used a broad string ("OSINT" OR "Open Source Intelligence" OR "Open-Source Intelligence" OR "open source intelligence" OR "open-source intelligence"), limited to publications from 2020 to 2023 to capture recent advancements, spanning four databases: Web of Science, IEEE Xplore, ACM Digital Library, and arXiv. A total of 60 studies were analyzed using the SQR3 method (Robinson, 1970) and organized into a structured Excel spreadsheet for each IC phase.

Technologies were categorized and analyzed for interrelationships within the OSINT framework. Verification was done using a Python script that scanned the papers for predefined keywords. The verified categories and relationships informed the artifact's development based on validated concept matrices.

For this study, a TR was selected as the artifact. This strategic tool identifies, monitors, and evaluates trends affecting industries or organizations, typically using circular diagrams with layers representing relevance or impact (Wulf et al., 2017).

3.4. Evaluation of the Design Specifications

The IC forms the foundation of the TR, offering clarity and an intuitive framework that simplifies technology extraction and research gap identification. Its design also ensures applicability across diverse intelligence disciplines in Germany, where this research was conducted. By emulating the structure of the German Federal government's TR (Stich et al., 2022),

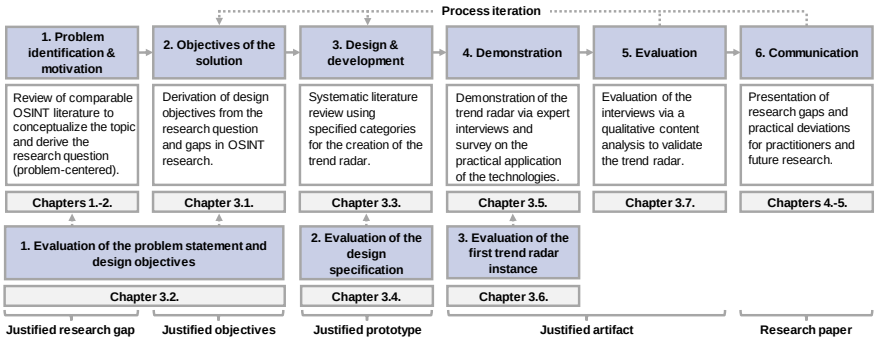

Figure 2. Iterative DSR model (Peffers et al., 2007) with control steps (Sonnenberg and vom Brocke, 2012)

the categories included in this radar have been limited to use case, technology, and maturity level. This focus enhances robustness, user-friendliness, and appropriate detail in the artifact. The concept matrices enable regular updates to the radar, ensuring current and standardized design. Rigorous verification of internal consistency maintains categorization integrity, meeting critical evaluation criteria in contemporary design research (Sonnenberg and vom Brocke, 2012).

3.5. Demonstration

The TR was demonstrated via guideline-based, systematizing expert interviews. Particularly in less structured and sparsely linked subject areas, this method enables dense data collection (Meuser and Nagel, 1991), especially when access to the social field is limited (Gläser and Laudel, 2009).

Experts (Table 1) were selected using theoretical sampling (Glaser and Strauss, 1967), focusing on Germany, the study's primary country of interest, to evaluate relevant trends. At least one expert from a security authority, the security industry, and a startup was chosen to capture diverse perspectives. A 'prestigious' company position ensures respondents possess relevant research knowledge.

The qualitative data collection utilized semi-structured interviews, uncovering underlying theoretical relationships (Bogner et al., 2014). The interview guide, based on the IC, commenced with a presentation of the TR. First, open questions were posed to compare it with respondents' practical experience, reducing subjectivity. Exploratory questions guided conversation flow, followed by specific closed questions for targeted follow-up (Saunders et al., 2012). The interview guide was pilot-tested with an expert. The interviews, contucted online, lasted up to an hour, with three main questions per phase (Bogner et al., 2014).

3.6. Evaluation of the First Instance of the Trend Radar

The TR demonstration confirmed its intuitive usability and usefulness in outlining OSINT technologies. Practitioners confirmed its completeness and consistency (cf. E1; E3; E4). The radar proved suitable for identifying research gaps and guiding practitioners, meeting the essential evaluation criteria (Sonnenberg and vom Brocke, 2012).

3.7. Evaluation

The evaluation employed qualitative data analysis to extract, synthesize, and structure interview data using a predefined search grid, enabling targeted summarization of relevant cross-interview information via a 'top-down approach' (Bogner et al., 2014).

Interviews were transcribed and analyzed with MAXQDA software for qualitative analysis. The categorization grid was developed, with first-level categories matching the Intelligence Cycle phases. Second-level categories reflect expert support or contradiction of the theory, while third-level categories indicate identified use cases. A 'general statements' category was included for overarching remarks, and fourth-level categories classify individual technologies. In total, 257 statements were categorized.

4. Results

The TR (Figure 3, 4) is read from the outside edge to the inside core. Each one-fifth of the cycle represents an IC phase. Subdivisions indicate phase-specific use cases, while color gradations show maturity levels. Numbered black and white dots denote grouped technologies, presented in a boxplot-like format reflecting varying maturity levels.

Table 1. Interviewed experts

ID	Organization	Position	Interview date
E1	Industry/ Authority	Senior Intelligence Consultant	07-14-2023
E2	Industry/ Authority	Referent Corporate Security	07-19-2023
E3	Authority	In-House Senior Consultant	07-28-2023
E4	Start-up	Managing Director of a German start-up	08-02-2023

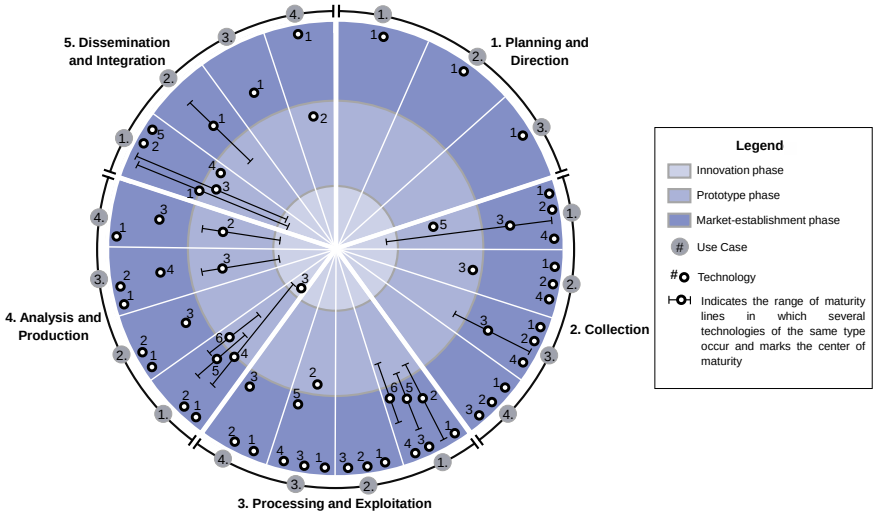

Figure 3. Resulting trend radar based on the Intelligence Cycle

4.1. Intelligence Cycle in Theory-Practice Comparison

Studies align with each phase of the IC but no application covers all phases as a third-generation OSINT tool. The literature mainly focuses on the collection phase, followed by analysis/production, and processing/exploitation. The dissemination/integration phase is least covered, followed closely by planning/direction. These findings align with the experts' practical experiences. They regard the IC as "state of the art" (cf. E3) but note different manifestations of the phases in practice (cf. E4). The planning/direction phase is often neglected, despite its crucial importance, leading to wasteful production (cf. E3). Conversely, OSINT is frequently associated solely with the collection phase, resulting in subpar outcomes due to high volumes of low-quality data (cf. E1; E2; E3). The main reason for this is that the IC is operated by at least three groups of people. Firstly, the customers, usually located at the "decision-maker level", with a primarily legal professional background (cf. E2). The second is the technician who carries out the data collection and processing (cf. E2; E3). Lastly, the analyst evaluates the data and creates the intelligence product (cf. E1). The process thereby is rarely transparent between the parties (cf. E1; E2) and is rarely anchored at the organizational level (cf. E4). According to the experts, there is thus no third-generation OSINT tool in use, at least not in German authorities. In addition, the collection focus is driven by concerns about missing vital information, later being revealed as publicly available (cf. E1).

4.2. Use Cases in Theory-Practice Comparison

Five main use cases emerged from the research: cyber security, health, security, journalism, and competition analysis. Cyber security studies primarily focus on Open Source Cyber Threat Intelligence (OSCTI), which involves collecting, monitoring, and analyzing public data to detect potential cyber threats (Ahuja et al., 2022). Health applications mainly address COVID-19 outbreak investigations (Kpozehouen et al., 2020). The security use case includes applications such as analyzing violent behavior in public transport (Nobili et al., 2021). The identified journalism study examines the Twitter activities of the OSINT journalists' association 'Bellingcat' (Bär et al., 2023). Competitive analysis involves, e.g., the performance classification of Chinese logistics companies (Tao et al., 2023).

1	**Planning and Direction**	2.3.3	Web crawler and/or web scraper	3.3.4	Statistical methods	4.3.4	Machine Learning
1.1	**Cyber security**	2.3.4	Other open source tools	3.3.5	Deep Learning	**4.4**	**Journalism, competitive analysis, general approach**
1.1.1	Manual	**2.4**	**Journalism, competitive analysis, general approach**	**3.4**	**Journalism, competitive analysis, general approach**	4.4.1	Manual
1.2	**Health**	2.4.1	Manual	3.4.1	Manual	4.4.2	Deep Learning
1.2.1	Manual	2.4.2	API interface	3.4.2	Standardized methods/algorithms/tools	4.4.3	Machine Learning
1.3	**Security**	2.4.3	Web crawler and/or web scraper	3.4.3	Natural Language Processing filters and labeling methods/tools	**5**	**Dissemination and Integration**
1.3.1	Manual	**3**	**Processing and Exploitation**	**4**	**Analysis and Production**	**5.1**	**Cyber security**
2	**Collection**	**3.1**	**Cyber security**	**4.1**	**Cyber security**	5.1.1	Files/reports
2.1	**Cyber security**	3.1.1	Manual	4.1.1	Manual	5.1.2	Dashboard/visualization map
2.1.1	Manual	3.1.2	Keyword/dictionary/hashtag filter	4.1.2	Standardized methods/algorithms and tools	5.1.3	Web interface/web application/online platform
2.1.2	API interface	3.1.3	Standardized methods/algorithms/tools	4.1.3	Tool stack	5.1.4	Automated alerts
2.1.3	Web crawler and/or web scraper	3.1.4	Statistical methods	4.1.4	Deep Learning	5.1.5	Graph creation
2.1.4	Other open source tools	3.1.5	Natural Language Processing filters and labeling methods/tools	4.1.5	Machine Learning	**5.2**	**Health**
2.1.5	Web application	3.1.6	Deep Learning	4.1.6	Artificial Intelligence	5.2.1	Web interface/web application/online platform
2.2	**Health**	**3.2**	**Health**	**4.2**	**Health**	**5.3**	**Security**
2.2.1	Manual	3.2.1	Manual	4.2.1	Manual	5.3.1	Dashboard/visualization map
2.2.2	API interface	3.2.2	Keyword/dictionary/hashtag filter	4.2.2	Statistical methods	**5.4**	**Competitive analysis, general approach**
2.2.3	Web crawler and web scraper	3.2.3	Standardized methods/algorithms/tools	4.2.3	Machine Learning	5.4.1	Dashboard/visualization map
2.2.4	Other open source tools	**3.3**	**Security**	**4.3**	**Security**	5.4.2	Web interface/web application/online platform
2.3	**Security**	3.3.1	Manual	4.3.1	Standardized methods/algorithms and tools		
2.3.1	Manual	3.3.2	Keyword/dictionary/hashtag filter	4.3.2	Statistical methods		
2.3.2	API interface	3.3.3	Standardized methods/algorithms/tools	4.3.3	Deep Learning		

Figure 4. Content of trend radar categorized by Intelligence Cycle phases, use cases, and technologies

Additionally, two identified studies focus generally on creating knowledge graphs on OSINF (Open Source Information, Hu et al., 2023; Ma et al., 2022).

The Experts note that OSINT is used in all authorities and various use cases, even if not explicitly labeled (cf. E2). It is most commonly applied in (cyber) security and OSCTI, especially within German Armed Forces, (Federal) Intelligence Service, domestic intelligence, and police (cf. E1; E2).

4.3. Technologies and Maturity Levels in Theory-Practice Comparison

Automated technologies are used in all phases and use cases except the initial one. While these technologies exhibit considerable market maturity, manual activities remain prevalent. Particularly cyber security shows the highest level of automation.

The most advanced automated technologies in the collection phase are web crawlers and scrapers. Established tools include 'off-the-shelf' options (cf. Middleton et al., 2020) and open-source solutions like 'Tweepy', a Python library for Twitter crawlers (e.g., Adewopo et al., 2020). More advanced prototypes combine parallelized, recursive, source-specific web crawlers and scrapers for enhanced data collection (e.g., Jenkins et al., 2021). A prototype method called 'focused crawling' adapts the crawling path dynamically using a content-driven ML algorithm, BERT ('Bidirectional Encoder Representation from Transformers', Kuehn et al., 2023). Technologies for crawling the dark web, like 'Torsion' (Sonawane et al., 2022), belong to the innovation phase. Experts note an increasing use of open-source tools alongside manual work (cf. E1; E3). However, they find traditional web crawling and scraping outdated due to errors and implementation challenges. Screenshot-based 'web shooting' with OCR ('Optical Character Recognition') extraction is seen as more modern and robust (cf. E3).

NLP ('Natural Language Processing') methods such as 'topic classification', 'part-of-speech tagging', and 'entity and relation annotation' demonstrate high automation levels in the processing/exploitation phase. Common technologies include the 'Python NLTK Toolkit' (Hubbard et al., 2022) and the 'Stanford CoreNLP Toolkit' (Middleton et al., 2020). Additionally, 'Deep Learning' (DL), particularly through 'word embedding' with the 'word2vec' algorithm, is prominent (e.g., Bai et al., 2020). Experts note that this phase mainly involves manual work within authorities due to the need for domain knowledge (cf. E1; E2; E3). The degree of automation varies with task abstraction: operational tasks requiring specific information show lower automation than long-term strategic analyses needing extensive data (cf. E3).

The highest automation level is seen in the analysis/production phase, where AI, ML, and DL technologies are prevalent. Under DL, vectorization algorithms, especially BERT versions, are used (e.g., Ma et al., 2022). ML models like BERT and 'Supervised Support Vector Machines' (e.g., Iorga et al., 2020) are also included, along with 'Random Forests', 'XGBoost', 'lightGBM', 'Naive Bayes', and 'Logistic Regression'. Publications frequently utilize multiple

algorithms for performance comparison (e.g., Tao et al., 2023) or layered analysis (e.g., Yang et al., 2022). AI technologies are less specified, except for Dale et al. (2023), who developed a bidirectional recurrent neural network with BiGur ('Bidirectional Gated Recurrent Unit') layers. Using modular public models, technologies are mainly classified as market-ready. The Experts indicate that this phase largely relies on manual content analysis due to limited technological understanding and acceptance in German authorities (cf. E2; E3; E4). Ethical and legal barriers, like GDPR ('General Data Protection Regulation'), hinder technology adoption (cf. E2; E4). Security concerns favor standalone systems, with smart technologies often used unofficially (cf. E2; E4). However, there is a need for modular, expandable systems to keep pace with advancements (cf. E1; E3; E4). Nonetheless, human experience and specialization are crucial for product quality, while the potential of 'Large Language Models' (LLM) remains uncertain (cf. E4).

In the dissemination/integration phase, tools like 'Power BI' (Tao et al., 2023) are used to create dashboards and visualizations. Interfaces, including Python GUIs and browser applications (Elmas et al., 2022), along with input masks for entire tool stacks (Arjun et al., 2020) are developed. Automated alert technologies for cyber security risk assessments are also common (Ahuja et al., 2022), and graph-based visualizations utilize tools/libraries like 'Matplot', 'Networkx', 'Pygraphistry', or the 'Neo4j-Browser' (Middleton et al., 2020). Except for alerts, the retrieval of results is largely semi-automated, with technologies in the market establishment phase. No information on user tests or new development involving user feedback was found in any studies. Experts state that automation within authorities during this phase remains very limited. The final product often consists of only a PDF document, email, or verbal report (cf. E1), which suffices in many cases (cf. E3). However, various automated tools beyond OSINT exist that could be applied (cf. E4), and there is a lack of necessary feedback for product improvement in practice (cf. E3).

5. Discussion and Conclusion

5.1. Contributions and Implications

The investigation into the existence of a robust, automated third-generation OSINT system (e.g., Ghioni et al., 2023) concludes negatively for Germany. Identified applications do not fully cover the IC, particularly lacking in the planning/direction and dissemination/integration phases, making human analysis essential. Numerous intelligent tools were identified in other phases, but integration has not met theoretical potential. This finding does not support the thesis of Yogish and Krishna (2021) that automated, AI-driven solutions are indispensable in all OSINT domains. The key research question is why proven applications have not gained widespread use, especially in German intelligence authorities. Future research should also explore enhancing technical support for the initial and final phases of the IC.

These questions require resolving research gaps (RGs) across the three key IC groups:

RG1: There is a gap in tools for the initial phase of the IC, in frameworks for requirements definition and communication (Section 4.1, cf. E3).

RG2: Effective dissemination and integration mechanisms tailored for authorities are lacking, primarily due to inadequate user testing and feedback incorporation. Established frameworks emphasize the importance of consumer feedback for product quality and data overload mitigation (NATO, 2001; USJFCOM, 2013).

RG3: The future of OSINT systems relies on modular concepts, yet research in this area is limited (Arjun et al., 2020; Wright et al., 2020).

RG4: It is crucial to move away from monolithic stand-alone setups in procurement procedures (Section 4.3, cf. E1; E3; E4).

RG5: Ensuring compliance with ethical and legal principles is vital for product adoption, necessitating robust legislative updates and adherence to regulations like GDPR (European Parliament, 2016; Wittmer and Platzer, 2022).

RG6: Addressing challenges requires a foundational technical understanding among decision-makers to foster openness to technology and enhance information sharing (NATO, 2001).

RG7: While LLMs show promise in intelligence analysis, their operational application remains underexplored (Radford et al., 2019).

RG8: Technicians often work independently in the IC, yet there is a lack of robust collection tools to match the rapidly evolving media landscape (Section 4.3, cf. E4).

RG9: Coordination gaps between analysts and technicians risk excessive data collection, showing a need for tools to improve transparency and mitigate biases (Lowenthal, 2020).

5.2. Limitations and Future Research

The first limitation is the lack of clarity regarding the legal and ethical basis (Ghioni et al., 2023; Wittmer and

Platzer, 2022), preventing verification of whether only public sources (NATO, 2002) were used in the analyzed studies. Additionally, it was not confirmed whether the technologies met the legal and ethical requirements for the use of the information obtained (Pastor-Galindo et al., 2020; Wittmer and Platzer, 2022).

The second limitation arises from classification categories not fully aligning with the MECE (Mutually Exclusive and Collectively Exhaustive) principle (Lee and Chen, 2018), particularly regarding the hierarchical dependency of AI, ML, and DL technologies. While the authors' wording was followed for objective reproduction, the accuracy of the information was not reviewed in detail. Furthermore, no fixed limits could be defined for the volume category, as these were not uniformly recorded in the studies.

The third limitation involves the small sample size of expert interviews, which were confined to Germany. Due to the difficulty in accessing this target group, interviews were not conducted with active users and decision-makers within authorities. Independent verification of coding by a second researcher is also recommended for improved intercoder reliability (Gläser and Laudel, 2009). Lastly, the research followed a linear execution rather than the suggested iterative approach (Peffers et al., 2007).

Despite these limitations, this study provides a comprehensive OSINT knowledge base through a practice-evaluated TR. It introduces a structured mechanism for capturing rapid developments and sheds light on the public security sector. Moreover, it serves as a guide for practitioners. Initial evaluations reveal two unanswered research questions and nine research gaps, highlighting critical areas for further exploration. Future research could apply this framework in other countries, where intelligence agencies may have more digitized processes, such as in allied nations (cf. E2). Extending the framework to other intelligence disciplines or domains, like the medical sector, could also offer new opportunities (cf. E4). The TR aims to elevate OSINT within academic research as an essential tool in today's complex world.

References

Adewopo, V., Gonen, B., & Adewopo, F. (2020). Exploring open source information for cyber threat intelligence. *IEEE Big Data*, 2232–2241.

Ahuja, K., Khushi, Dipali, & Sharma, N. (2022). Cyber security threats and their connection with twitter. *IEEE ICAIS*, 1458–1463.

AlKilani, H., & Qusef, A. (2021). Osint techniques integration with risk assessment iso/iec 27001. In J. A. Lara Torralbo, S. A. Aljawarneh, V. Radhakrishna, & A. N. (Eds.), *Int. conf. on data science, e-learning and information systems* (pp. 82–86). ACM.

Arjun, A. V., Buvanasri, A. K., Meenakshi, R., Karthika, S., & Ashok, K. M. (2020). Peoplexploit: A hybrid tool to collect public data. *IEEE ICCCSP*, 1–6.

Bai, C., Li, A., Gao, Z., & Cui, X. (2020). Research on anti-terrorism intelligence mining method based on attention neural networks. *IEEE ICCASIT*, 458–464.

Bär, D., Calderon, F., Lawlor, M., Lickleder, S., Totzauer, M., & Feuerriegel, S. (2023). Analyzing social media activities at bellingcat. *Proceedings of the 15th ACM Web Science Conference 2023*, 163–173.

Benes, L. (2013). Osint, new technologies, education: Expanding opportunities and threats. a new paradigm. *Journal of Strategic Security*, *6*(3Suppl), 22–37.

Billings, C. E. (1997). *Aviation automation: The search for a human-centered approach*. Lawrence Erlbaum Associates Publishers.

Bleck, S. (2004). *Entwicklung einer Methodik zur integrierten Planung von Informationstechnologie-Einsatz und intermediären Informationsdienstleistungen im elektronischen Geschäftsverkehr* (Vol. 72). Shaker.

Bogner, A., Littig, B., & Menz, W. (2014). *Interviews mit Experten: Eine praxisorientierte Einführung*. Springer.

Böhm, I., & Lolagar, S. (2021). Open source intelligence. *International Cybersecurity Law Review*, *2*(2), 317–337.

Central Intelligence Agency. (1987). *Factbook on intelligence* [Retrieved April 15, 2023, from https://fh.ms/osint-factbook].

Central Intelligence Agency. (2023). *The intelligence cycle: Briefing* [Retrieved April 15, 2023, from https://fh.ms/osint-icbriefing].

Cleven, A., Niehaves, B., Plattfaut, R., Riemer, K., Simons, A., & vom Brocke, Jan, Martin Hilti. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. *17th European Conference on Information Systems*.

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, *1*(1), 104–126.

Dale, D., McClanahan, K., & Li, Q. (2023). Ai-based cyber event osint via twitter data. *IEEE ICNC*, 436–442.

Delen, D., & Demirkan, H. (2013). Data, information and analytics as services. *Decision Support Systems*, *55*(1), 359–363.

Dokman, T., & Ivanjko, T. (2020). Open Source Intelligence (OSINT): issues and trends. *INFuture2019*.

Dos Passos, D. S. (2017). Big data, data science and their contributions to the development of the use of open source intelligence. *Sistemas & Gestão*, *11*(4), 392–396.

Duncheon, C. (2002). Product miniaturization requires automation – but with a strategy. *Assembly Automation*, *22*(1), 16–20.

Elmas, T., Ibanez, T. R., Hutter, A., Overdorf, R., & Aberer, K. (2022). Waypop machine: A wayback machine to investigate popularity and root out trolls. *2022 IEEE/ACM ASONAM*, 391–395.

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462–492.

European Parliament (Ed.). (2016). Document 32016R0679: Regulation (EU) 2016/679 (General Data Protection Regulation).

García Lozano, M., Brynielsson, J., Franke, U., Rosell, M., Tjörnhammar, E., Varga, S., & Vlassov, V. (2020). Veracity assessment of online data. *Decision Support Systems*, *129*, 113132.

Ghioni, R., Taddeo, M., & Floridi, L. (2023). Open source intelligence and ai: A systematic review of the gelsi literature. *AI & Society*, 1–16.

Gibson, H. (2016). Acquisition and preparation of data for osint investigations. In B. Akhgar, P. S. Bayerl, & F. Sampson (Eds.), *Open source intelligence investigation* (pp. 69–93). Springer.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Gläser, J., & Laudel, G. (2009). *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*. VS Verlag.

Herrera-Cubides, J. F., Gaona-García, P. A., & Sánchez-Alonso, S. (2020). Open-source intelligence educational resources: A visual perspective analysis. *Applied Sciences*, *10*(21), 7617.

Hu, Y., He, L., Tang, X., Luo, G., He, S., & Fang, Q. (2023). Construction of domain knowledge graph based on open source intelligence. *2023 IEEE 2nd EEBDA*, 1378–1382.

Hubbard, J., Bendiab, G., & Shiaeles, S. (2022). Ipass: A novel open-source intelligence password scoring system. *IEEE CSR*, 90–95.

Hwang, Y.-W., Lee, I.-Y., Kim, H., Lee, H., & Kim, D. (2022). Current status and security trend of osint. *Wireless Communications and Mobile Computing*, *2022*, 1–14.

Iorga, D., Corlatescu, D., Grigorescu, O., Sandescu, C., Dascalu, M., & Rughinis, R. (2020). Early detection of vulnerabilities from news websites using machine learning models. *2020 19th RoEduNet Conference*, 1–6.

Ish, D., Ettinger, J., & Ferris, C. (2022). *Evaluating the effectiveness of artificial intelligence systems in intelligence analysis* (Vol. RR-A464-1). RAND Corporation.

Jenkins, D., Liebrock, L. M., & Urias, V. (2021). Designing a modular and distributed web crawler focused on unstructured cybersecurity intelligence. *2021 ICCST*, 1–6.

Kayser, F. (2024). Addressing Challenges in a Dangerous World: Developing a Design Science Artifact for Advancing Open Source Intelligence (OSINT) Research [https : //fh.ms/osint-paper].

Kpozehouen, E. B., Chen, X., Zhu, M., & Macintyre, C. R. (2020). Using Open-Source Intelligence to Detect Early Signals of COVID-19 in China: Descriptive Study. *JMIR public health and surveillance*, *6*(3), e18939.

Kuehn, P., Schmidt, M., Bayer, M., & Reuter, C. (2023). ThreatCrawl: A BERT-based Focused Crawler for the Cybersecurity Domain (PEASEC, Technical University of Darmstadt, Ed.). *arXiv*.

Lee, C.-Y., & Chen, B.-S. (2018). Mutually-exclusive-and-collectively-exhaustive feature selection scheme. *Applied Soft Computing*, *68*, 961–971.

Lowenthal, M. M. (2020). *Intelligence: From secrets to policy* (8th edition). Sage.

Ma, H., Liu, X., & Zhao, W. (2022). Research on domain-specific knowledge graph based on the roberta-wwm-ext pretraining model. *Computational intelligence and neuroscience*, *2022*.

Meuser, M., & Nagel, U. (1991). ExpertInneninterviews - vielfach erprobt, wenig bedacht: ein Beitrag zur qualitativen Methodendiskussion, 441–471.

Middleton, S., Lavorgna, A., Neumann, G., & Whitehead, D. (2020). Information extraction

from the long tail. *12th ACM Conference on Web Science Companion*, 82–88.

Nobili, M., Faramondi, L., Setola, R., Ghelli, M., Persechino, B., & Lombardi, M. (2021). An osint platform to analyse violence against workers in public trasportation. *IEEE ICCSI*, 1–6.

North Atlantic Treaty Organization. (2001). *NATO Open Source Intelligence Handbook* [Retrieved April 21, 2023, from https://fh.ms/osint-nato].

North Atlantic Treaty Organization. (2002). *NATO Open Source Intelligence Reader* [Retrieved April 21, 2023, from https://fh.ms/osint-natoreader].

Pastor-Galindo, J., Nespoli, P., Gomez Marmol, F., & Martinez Perez, G. (2019). Osint is the next internet goldmine: Spain as an unexplored territory. *5th Nat. Conf. Cybersecur. (JNIC)*.

Pastor-Galindo, J., Nespoli, P., Gomez Marmol, F., & Martinez Perez, G. (2020). The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends. *IEEE Access*, *8*, 10282–10304.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45–77.

Phythian, M. (Ed.). (2013). *Understanding the intelligence cycle*. Routledge.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI*.

Reuser, A. (2017). The ris open source intelligence cycle. *Journal of Mediterranean and Balkan Intelligence*, *10*(2).

Robinson, F. P. (1970). *Effective study* (4th ed.). Harper & Row.

Saunders, M., Lewis, P., & Thornhill, A. (2012). *Research methods for business students* (6. ed.). Pearson.

Singh, S., & Singh, N. (2012). Big data analytics. *Int. Conf. on Communication, Information & Computing Technology*, 1–4.

Smith-Boyle, V. (2022). How osint has shaped the war in ukraine [Retrieved April 15, 2023, from https://fh.ms/osint-boyle]. *American Security Project (ASP)*.

Sonawane, H. S., Deshmukh, S., Joy, V., & Hadsul, D. (2022). Torsion: Web reconnaissance using open source intelligence. *IEEE CONIT*, 1–4.

Sonnenberg, C., & vom Brocke, J. (2012). Evaluations in the science of the artificial – reconsidering the build-evaluate pattern in design science research. In *Design science research in information* (pp. 381–397, Vol. 7286).

Stich, V., Stroh, M.-F., Abbas, M., Frings, K., & Kremer, S. (2022). *Digitalisierung der Wirtschaft in Deutschland: Technologie- und Trendradar 2022* [Retrieved April 15, 2023, from https://fh.ms/osint-trendradar]. BMWK.

Tao, Z., Charoenkwan, P., Paphawasit, B., & Rujeerapaiboon, N. (2023). Machine learning-based classification of competitors performance: Evidence from chinese logistics companies. *IEEE ECTI DAMT & NCON*, 131–137.

U.S. Joint Force Command. (2013). Joint intelligence. In *Joint Publication 2-0 (JP 2-0)*.

vom Brocke, J., Hevner, A., & Maedche, A. (2020). Introduction to design science research. In J. vom Brocke, A. Hevner, & A. Maedche (Eds.), *Design science research. cases* (pp. 1–13). Springer.

Webster, J., Watson, & T. Richard. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*, xiii–xxiii.

Williams, H. J., & Blum, I. (2018). *Defining second generation open source intelligence (osint) for the defense enterprise*. RAND Corporation.

Wittmer, S., & Platzer, F. (2022). Zulässigkeit von Open Source-Ermittlungen zur Strafverfolgung im Darknet. In *INFORMATIK (LNI)*. Gesellschaft für Informatik, Bonn.

Wright, T., Whitfield, S., Cahill, S., & Duffy, J. (2020). Project umbra. *2020 IEEE/ACM Int. Conf. ASONAM*, 748–751.

Wulf, J., Mettler, T., & Brenner, W. (2017). Using a trend radar for competitive intelligence: Lessons from an industrial product service system manufacturer. *Technological Forecasting and Social Change*, *127*, 118–130.

Yang, J.-Z., Liu, F., Zhao, Y.-J., Liang, L.-L., & Qi, J.-Y. (2022). NiNSRAPM: An ensemble learning based non-intrusive network security risk assessment prediction model. *7th IEEE DSC*, 17–23.

Yogish, P. U., & Krishna, P. K. (2021). Open source intelligence and its applications in next generation cyber security - a literature review. *International Journal of Applied Engineering and Management Letters*, *5*(2).