31/3/2023

Stato dell'arte in ambito OSINT/SIEM, con valutazione e motivazione per piattaforma OSINT/SIEM scelta come riferimento

D.2.1.1 Fighting Cybercrime with OSINT

M. Gnaldi, L. Grilli, A. Milani, A. Navarra, M.C. Pinotti, V. Poggioni, F. Santini, C.Taticchi,
UNIVERSITÀ DEGLI STUDI DI PERUGIA

Sommario

Tabella delle Figure

Introduzione e terminologia riguardante le piattaforme SIEM

In questo deliverable di progetto (D.2.1.1, Progetto di Ateneo FICO), ci concentriamo sulla descrizione dello stato dell'arte relativo ai sistemi di **Security Information and Event Management** (**SIEM**) che fanno parte di sistemi di **Security Operations and Incident Management** (**SOIM**). Le origini dei SOIM possono essere individuate a partire dal 1981 nel documento di James Anderson [1].

I SOIM possono essere visti come un'applicazione e un'automazione del ciclo di **Monitor Analyze Plan Execute-Knowledge** (**MAPE-K**) per la sicurezza informatica [2], anche se questo modello è stato definito più tardi rispetto agli sviluppi iniziali dei sistemi SOIM. La computazione deve essere il più possibile autonoma, in modo ad adattare i sistemi ICT alle condizioni operative, che possono risultare alquanto mutevoli in una situazione di attacco. Il ciclo MAPE-K, che è descritto in Figura 1, è guidato da eventi che forniscono informazioni sul comportamento corrente del sistema. I vari passaggi sequenziali del ciclo analizzano il flusso di eventi (*traccia*) per fornire feedback al sistema, modificandone il comportamento in base a osservazioni e policy predefinite, e consentendo inoltre l'adattamento automatico per fornire al meglio il servizio agli utenti. I sistemi SOIM hanno visto un innalzamento in termini di automazione e complessità nel corso degli anni, come risultato della crescente necessità a fornire i servizi dell'infrastruttura ICT in maniera appropriata. Questi aggiornamenti e miglioramenti hanno mano a mano interessato la maggior parte dello spettro del ciclo MAPE-K. Il dominio SOIM suppone quindi che il flusso di lavoro realizzato dal ciclo MAPE-K sia implementato in componenti tecnici-tecnologici presenti all'interno della struttura ICT da proteggere.

Nella attività di *monitoring* (*monitor* in Figura 1), ci si concentra sulle varie sorgenti di dati in modo da controllare ciò che interessa. Già nella fase di gli algoritmi di *analyze*, gli algoritmi di *detection* cercano di determinare se l'informazione acquisita nella fase precedente costituisce evidenza di attacco. L'attività di *plan* (sempre Figura 1) fornisce un insieme di contromisure per rispondere all'attacco, mentre nella attività finale di *execute*, queste contromisure vengono eseguite.

La Figura 1 mostra le funzionalità e come interagiscono i diversi componenti che fanno parte del flusso generale di lavoro di un SOIM. Gli **Intrusion Detection Systems** (**IDS**) sono la prima componente di un SOIM studiata in ordine temporale (prime commercializzazioni nella seconda metà degli anni '90), e coprono le fasi *monitoring* e *detection*. Queste due componenti sono però insufficienti a causa della necessità di processare gli eventi in tempo reale e dalla limitata copertura fornita dalla singola fonte di dati su cui lavora e si effettua detection. La seconda componente è quindi rappresentata dai SIEM, che estendono l'attività di analisi fornendo una visione di insieme da più sorgenti di dati componendole insieme, filtrandole e aggregando la grande quantità di *alert* proveniente dalle attività di detection sulle singole fonti. L'attività di plan (non implementata dagli IDS) è in sostanza il cuore di un SIEM, al fine di produrre un piano di risposta all'attacco e da eseguire solitamente attraverso processi manuali da parte degli operatori. Le piattaforme SIEM di grandi dimensioni, complesse e costose, trovano loro collocazione nei **Security Operating Center** (**SOC**), che forniscono sia risorse sia tecnologiche che umane. Più recentemente, i sistemi **Security Orchestration, Automation and Reponse** (**SOAR**) hanno portato la fase di plan e execute al fine di rendere possibile una risposta automatizzata all'attacco.

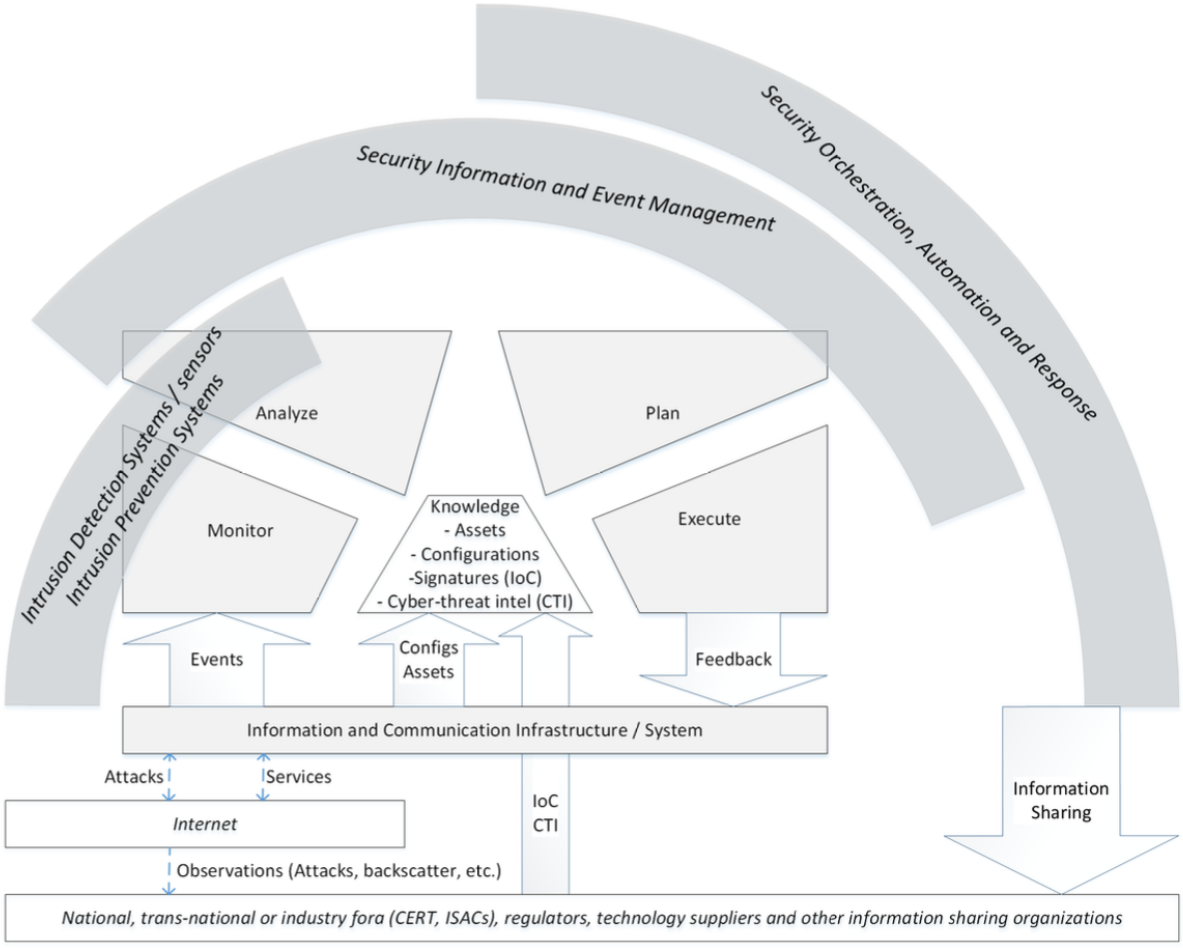

Figura 1: Ciclo di computazione autonoma MAPE-K [3]

SIEM

Dal punto di vista dell'analisi (attività Analyze), un SIEM mira a fornire ulteriori informazioni sulle attività sospette segnalata dai sensori monitorati. A causa del volume degli eventi e della natura in tempo reale del rilevamento eseguito dagli IDS, questi sensori di solito ispezionano una singola fonte di informazioni in una posizione specifica dell'infrastruttura ICT. Pertanto, è difficile per loro rilevare

attacchi su larga scala o distribuiti. Quindi, la centralizzazione degli alert, che è la prima caratteristica centrale delle piattaforme SIEM, consente l'utilizzo di ulteriori algoritmi di detection che possono indicare anomalie e attacchi non sufficientemente segnalati dai singoli sensori. L'anomalia può però diventare più significativa quando le informazioni vengono invece aggregate.
I SIEM inoltre contribuiscono in modo sostanziale all'attività di plan in un SOIM (Figura 1). Essi si possono considerare come un sistema di supporto alle decisioni (**decision support system**). Per quanto riguarda l'attività di plan, le piattaforme SIEM hanno lo scopo di definire un insieme di azioni che possono essere intraprese per bloccare un attacco o mitigare i suoi effetti.
I fondamenti dei SIEM possono essere fatti risalire al Dicembre 1998, in un incontro organizzato dal *Defense Advanced Research Projects Agency* (*DARPA*). L'obiettivo originale era quello di consentire un confronto delle prestazioni dei vari progetti di ricerca sul rilevamento delle intrusioni che DARPA stava finanziando, e questo ha prodotto diversi lavori: il per esempio Lincoln Labs/KDD dataset [4] [5], ed in seguito tre differenti *Request for Comment* (RFC) che hanno fornito una prima formalizzazione dal punto di vista tecnico dei SIEM: i requisiti per il formato e lo scambio di informazioni di interesse per un SIEM in RFC 4766 [6], il formato dei messaggi di alert in RFC 4765 [7] (**Intrusion Detection Message Exchange Format** o **IDMEF**), ed infine in RFC 4767 [8] (**Intrusion Detection eXchange Protocol** o **IDXP**).

Collezione di dati

Il primo obiettivo di una piattaforma SIEM è quello di raccogliere e centralizzare le informazioni provenienti da più sensori. Diversi problemi devono essere risolti per fare in modo che ciò accada.
Prima di tutto, deve esserci un canale di comunicazione tra i sensori che forniscono gli avvisi e la piattaforma SIEM. Questo canale di comunicazione deve essere fortemente protetto, perché negli avvisi possono essere incluse informazioni sensibili. Deve anche essere adeguatamente dimensionato in modo che vi sia larghezza di banda sufficiente per trasportare le informazioni richieste. Poiché i sensori hanno spesso capacità di archiviazione limitate, la disponibilità del collegamento è essenziale.
In secondo luogo, il SIEM deve essere in grado di interpretare in modo coerente le informazioni fornite dai sensori. Data l'ampia gamma di fonti di dati e metodi di rilevamento disponibili, ciò richiede molto lavoro per abbinare le informazioni dagli avvisi con i formati di dati interni SIEM. L'approccio generale di una piattaforma SIEM consiste nel definire un'unica struttura dati per gli avvisi, spesso una singola tabella di database. Ciò significa che il database contiene molte colonne, ma che l'inserimento di un avviso comporta spesso un riempimento di solo alcune delle colonne (quelle specifiche per quella fonte di dati).
La raccolta dei dati è generalmente gestita dalla piattaforma SIEM, beneficiando delle estensioni dei sensori. I fornitori di piattaforme SIEM generalmente definiscono i propri connettori e formati, gestendo contemporaneamente sia il problema della sicurezza del trasporto che quello dell'importazione dei dati.
La comunicazione di un messaggio di alert riguarda tre livelli differenti: 1) lo *schema* definisce la struttura dei messaggi, il tipo e il significato degli attributi. Diversi schemi per alert si appoggiano alla struttura che *Common Vulnerabilities and Exposures* (*CVE*)[1] utilizzata per documentare gli attacchi. 2) la *codifica* definisce come tali messaggi e gli attutiscono codificati per formare stringhe di bit; alcuni esempi di formati testuali sono per esempio *Syslog*, *JSON*, *XML* o *YAML*, mentre esempi di codifica binaria sono *BER*, *CER* o *BSON*. Mentre i primi sono leggibili anche da un operatore umano, i secondi sono più compatti e quindi facilmente trasportabili. Infine, 3) i protocolli di trasporto descrivono come l'informazione codificata sia comunicata dal sensore al SIEM; esempi di

[1] CVE: https://cve.mitre.org

protocolli includono *Syslog*, *IDXP*, *HTTP* o *AMQP*. A quest'ultimo livello tipicamente vengono affidati il controllo degli accessi, la confidenzialità, la compressione, e l'affidabilità della comunicazione.
Di seguito riportiamo i formati di dato più famosi utilizzati in ambito SIEM.
Syslog (RFC 5424) [9] è lo standard di fatto per l'acquisizione degli alert da parte delle piattaforme SIEM. Esso è infatti ampiamente diffuso, facile da comprendere e analizzare, e piuttosto affidabile anche se, in caso si utilizzi UDP come protocollo a livello trasporto, non ci sia garanzia di consegna al destinatario. Nella pratica però, Syslog si è rivelato molto scalabile e ha quindi ottenuto successo nel suo utilizzo: è infatti ampiamente utilizzato dagli operatori di rete. Il suo svantaggio è la limitazione del suo schema (costituito da un timestamp, un'origine e una stringa di testo ASCII) e la scarsa dimensione massima del messaggio (limitata a 1000 byte).
Il **Common Event Format** (CEF) è il formato di scambio proprietario della piattaforma SIEM di Arcsight (una compagnia software deglis Stati Uniti).[2] È orientato alla definizione generale di eventi rilevanti per la sicurezza ed include le informazioni essenziali necessarie per descriverli. Questo formato è rappresentativo delle strutture "flat" utilizzate spesso nei database della piattaforma SIEM. Sebbene abbia un gran numero di attributi, alcuni non sono sufficientemente documentati per l'uso.
Il **Log Event Enhanced Format** (o LEEF) è il formato di scambio proprietario della piattaforma SIEM QRadar di IBM. Si concentra sugli eventi di sicurezza di rete e come tale non è così ricco come CEF.
Il **Common Information Model** (CIM) è uno standard della Distributed Management Task Force (DMTF).[3] È ampiamente utilizzato per la gestione di sistemi distribuiti. Poiché è molto generico, la sua espressività per gli eventi di sicurezza informatica è limitata.
Il formato **Cloud Auditing Data Federation** è ancora in fase di sviluppo, inizialmente come XDAS (XDAS/CADF), e è possibile una sua futura integrazione con DMTF. Esso si concentra su eventi di sistema e ambienti cloud.
Il formato **Common Event Expression** (CEE) è stato ideato da MITRE come formato standard per i file di log nei sistemi informatici. È stato sviluppato in collaborazione tra enti governativi statunitensi e fornitori SIEM. Esso separa chiaramente il formato del messaggio (CEE Event Model o Profile), la codifica (CEE Log Syntax) e il trasporto (CEE Log Transport). Sfortunatamente però, il lavoro su CEE si è interrotto.
L'**Intrusion Detection Message Exchange Format** (IDMEF) è un documento informativo dell'IETF. Non specifica uno standard e pertanto la sua adozione da parte dell'industria è stata molto limitata. È visto come un formato complesso, e infatti la sua specifica è di grandi dimensioni. Il formato IDMEF contiene un grande numero di attributi, il più grande di tutti i formati considerati. Il suo tentativo di essere molto esaustivo ha portato alcune delle sue strutture dati ad essere obsolete. La scelta dei messaggi codificato come XML crea anche un onere significativo nel trasporto.

L'ampio spettro delle specifiche riportate precedentemente dimostra che ad adesso non esiste consenso tra i fornitori SIEM e i fornitori di sensori per concordare il contenuto di un alert. Sebbene molte delle specifiche siano accessibili ai fornitori di sensori, i fornitori di piattaforme SIEM forniscono i loro connettori e si occupano di tradurre le informazioni del sensore nei propri formati, con il rischio di perdere informazioni o interpretare erroneamente il contenuto. La Figura 2 raccoglie le informazioni principali sui formati di rappresentazione di alert.

Formato	Owner	Trasporto	Codifica	Struttura	N. Attributi

[2] https://www.microfocus.com/en-us/cyberres/secops/arcsight-esm
[3] https://www.dmtf.org/

CEF	HP/Arcsight	Syslog	Key/value	Flat	117
LEEF	IBM/QRadar	Syslog	Key/value	Flat	50
CIM	DMTF	Any	XML	UML	58
CADF	The Open Group, DMTF, (NetIQ)	Any	JSON	Classi e attributi	48
CEE	MITRE	Syslog	JSON,XML	CEE Event Model e Profile	56
IDMEF	IETF	IDXP	XML	UML	166

Figura 2: Riassunto delle caratteristiche dei fari formati di rappresentazione di alert

Correlazione degli alert

La correlazione degli alert [13, 14], ha l'obiettivo di dare un senso al flusso di avvisi ricevuto dal SIEM piattaforma. La correlazione ha diversi obiettivi:

1. ridurre il numero di alert che l'analista deve elaborare raggruppando insieme gli avvisi;
2. aggiungere elementi contestuali per consentire un'analisi più accurata e rapida di un gruppo di alert;
3. aggiungere alert agli elementi di pianificazione e mitigazione di livello superiore in modo che vengano gestiti correttamente;
4. scartare gli alert considerati falsi positivi e che non richiedono ulteriore elaborazione.

È possibile individuare le seguenti tecniche di correlazione. La **correlazione tra alert** mira a raggruppare gli avvisi provenienti da uno o più sensori e che corrispondono alla stessa minaccia. I sensori IDS/IDPS tendono ad avere una visione ristretta del flusso di dati. Se gli eventi si verificano ripetutamente, ad esempio quando si propaga un malware all'interno della stessa rete, al SIEM verranno segnalati più alert. Raggruppare gli alert che corrispondono allo stesso fenomeno aiuta l'analista a riconoscere e a giudicarne l'importanza. La **correlazione tra allarmi e ambiente** è invece legata al contesto della rilevazione e cioè l'ambiente in cui si trovano i sensori. Le informazioni sull'ambiente provengono da molte fonti, le due più interessanti sono l'individuazione di tutti i dispostivi connessi alla rete ("network inventory") e le informazioni provenienti dalle scansioni di vulnerabilità. Queste due fonti identificano gli asset attivi e i rischi a cui sono potenzialmente soggetti. Questo tipo di correlazione è particolarmente interessante in quanto fornisce all'analista informazioni sull'impatto che stanno avendo gli alert. La **correlazione tra avvisi e fonti esterne** riguarda la *situational awareness*, la quale implica l'essere consapevoli di ciò che sta accadendo nelle vicinanze per capire come le informazioni, gli eventi e le proprie azioni avranno un impatto locale nel prossimo futuro. In questo modo è possibile comprendere in anticipo i percorsi che un utente malintenzionato potrebbe seguire, e aiutare quindi l'analista a decidere in modo proattivo di bloccare i progressi dell'attaccante, invece di reagire dopo l'evento. Infine, un'altra tendenza rilevante è lo scambio di informazioni riguardo incidenti di sicurezza. Attraverso pressioni

normative, gli operatori delle infrastrutture critiche sono tenuti a informare le autorità quando sono vittime di violazioni della sicurezza informatica (ad esempio, gli istituti bancari). La condivisione delle informazioni sulle violazioni aiuta le altre entità dello stesso dominio o che utilizzano tecnologie simili a proteggersi in modo proattivo.
L'approccio iniziale alla correlazione degli avvisi si è basato su regole. La correlazione basata su regole descrive esplicitamente le relazioni logiche tra avvisi o regole per inferire tali relazioni [14,15,16]. La prima generazione di piattaforme SIEM combinava database SQL fortemente strutturati e ad alte prestazioni con regole di interpretazione di motori logici. Questa prima generazione ha riscontrato due problemi: in primis il degrado delle prestazioni con l'aumento del volume degli avvisi, ed inoltre la difficoltà di creare e mantenere l'insieme delle regole. I database SQL comportano una significativa riduzione delle prestazioni per la procedura di indicizzazione. Mentre questa è utile per effettuare delle interrogazioni, le piattaforme SIEM sono caratterizzate da un'alta intensità di inserimento.
Nonostante l'aumento delle prestazioni e l'ottimizzazione del database, è stata sviluppata una seconda generazione di piattaforme SIEM sfruttando tecnologie di database NoSQL. Questo approccio big data, o ad alta intensità di dati, è iniziato molto presto utilizzando contatori [13], modelli statistici [17] o altre tecniche [18, 19]. Tecnologicamente, questo approccio può essere implementato tramite aggregazione di log e query di riepilogo, ad esempio grazie al noto stack ElasticSearch-Kibana-Logstash (ELK).[4] Questo approccio orientato ai dati è diventato molto comune oggi, in quanto è in grado di far fronte a grandi volumi di informazioni non strutturate in arrivo. La mancanza di una struttura relazionale potrebbe comunque introdurre incoerenze e confusione nei nomi, incidendo sulla capacità degli analisti di diagnosticare e mitigare le minacce. Inoltre, l'attenzione al volume potrebbe rendere difficile la gestione di fenomeni di attacco rari come gli APT.

Letteratura su piattaforme SIEM

L'analisi e la valutazione dei sistemi SIEM sono state ampiamente proposte in letteratura. Mentre alcune ricerche si concentrano sugli aspetti commerciali, altre si concentrano invece sulle caratteristiche tecniche che potrebbero essere migliorate nelle attuali soluzioni SIEM. Istituzioni ben note come Gartner, ad esempio, propongono un'analisi commerciale dei sistemi SIEM basata sul mercato e sui principali fornitori, per la quale viene pubblicato un rapporto su base annuale per individuare i fornitori SIEM leader di mercato, sfidanti, attori di nicchia, o visionari. Sebbene aziende come Gartner valutino periodicamente la capacità dei SIEM, non esiste un'indagine sistematica di questi sistemi, delle loro capacità e delle loro mancanze. Inoltre, altre istituzioni di sicurezza (ad esempio, Techtarget, Info-Tech Research Group), concentrano invece i loro report sulle capacità delle soluzioni SIEM e sul modo in cui i fornitori SIEM possono essere confrontati e valutati. Techtarget, da un lato, pubblica guide elettroniche periodiche sulla sicurezza dei sistemi SIEM e su come definire la strategia, la gestione e il successo SIEM nell'azienda.[5] Info-Tech, d'altra parte, fornisce rapporti tecnici sul panorama dei venditori SIEM concentrandosi sui vantaggi e gli svantaggi dei principali SIEM commerciali.[6] Entrambe le organizzazioni prendono il Gartner Magic Quadrant come punto di riferimento per la loro analisi, lasciando da parte gli aspetti più tecnici da considerare

[4] https://www.elastic.co/what-is/elk-stack
[5] https://www.techtarget.com/searchsecurity/essentialguide/How-to-define-SIEM-strategy-management-and-success-in-the-enterprise
[6] https://www.infotech.com/research/ss/it-vendor-landscape-plus-security-information-event-management

nei futuri SIEM. Allo stesso modo, organizzazioni come Solutions Review[7] offrono rapporti periodici per guidare gli acquirenti SIEM nella selezione appropriata della soluzione SIEM per le loro attività. Gli autori analizzano le funzionalità SIEM chiave ed eseguono una mappa dei fornitori di confronto basata su tre aspetti fondamentali: conformità, gestione dei log e rilevamento delle minacce. Sebbene il rapporto consenta di mettere in contatto i potenziali acquirenti con i fornitori, non fornisce dettagli tecnici degli strumenti né discute sulle potenziali capacità da migliorare negli attuali SIEM o sui fattori esterni che potrebbero influire sulle loro prestazioni in futuro.
Per quanto riguarda la letteratura scientifica invece, in [11] viene per esempio presentata un'analisi commerciale e tecnica di alcune delle principali soluzioni SIEM disponibili sul mercato, vale a dire ArcSight, QRadar, McAfee, LogRhythm, USM-OSSIM, RSA, Splunk e SolarWinds. Questa scelta si è basata sulle prestazioni e sul percorso commerciale delle aziende che hanno sviluppato questa tecnologia nell'ultimo decennio. Sebbene la maggior parte delle soluzioni analizzate fornisca interfacce grafiche di facile utilizzo, le capacità di visualizzazione e reazione sono giudicate come limitate per gestire un numero enorme di eventi raccolti. Un possibile miglioramento, secondo gli autori dell'articolo, sarebbe quindi quello di sviluppare estensioni di visualizzazione e analisi, che aiutino a fornire agli utenti un alto livello di comprensione della situazione e una capacità decisionale e di reazione più efficiente. Per quanto riguarda l'archiviazione dei dati e il prezzo, sebbene la maggior parte delle soluzioni analizzate includano buone capacità di archiviazione dei dati, sono limitate dalla disponibilità dell'hardware e di solito richiedono prodotti aggiuntivi (e licenze in base al volume di dati) con un conseguente aumento del prezzo. Le soluzioni sicure ed "elastiche" basate sullo storage cloud per l'archiviazione dei dati SIEM a lungo termine (ad es. Amazon S3, Amazon Glacier, Windows Azure, Blob Store, ecc.) sono viste come alternative promettenti.
In [12] gli autori presentano il comportamento e i miglioramenti di un SIEM nel contesto di un sistema *mission-critical* di una delle principali aziende leader nel dominio del controllo del traffico aereo. Tale sistema emette enormi volumi di log di testo (dati non strutturati). Gli autori presentano le sfide nell'affrontare tali file di log, il lavoro in corso di integrazione di un SIEM open source, ed infine le possibili direzioni nella modellazione delle linee di base di comportamento del sistema al fine di dedurre gli indicatori di compromissione.
Kotenko e Chechulin [20] propongono un framework per la modellazione degli attacchi e la valutazione della sicurezza nei sistemi SIEM applicabile per i futuri sistemi dell'Internet of Things. L'approccio si concentra sulle caratteristiche tecniche (ad esempio, la valutazione dell'utilizzo di un archivio di sicurezza interno, di un database di sicurezza aperto, grafici delle dipendenze dei servizi, grafici degli attacchi e metriche di sicurezza) da integrare in un framework SIEM al fine di migliorarne la funzionalità. Gli autori affermano di ottenere valutazioni più accurate e rapide degli aspetti di sicurezza della rete mediante l'uso del modello di attacco proposto. Oltre ad alcuni aspetti tecnici, non vengono prese in considerazione altre caratteristiche per il miglioramento degli attuali sistemi SIEM.
Poiché non tutte le aziende dispongono di strumenti, competenze e competenze interne per proteggere da sole gli ambienti Cloud, le soluzioni *Security as a Service* (*SecaaS*) stanno diventando sempre più popolari, promettendo risparmi sui costi e un'adeguata rilevazione e prevenzione delle minacce in tempo reale. In [21] gli autori delineano le attuali aree di ricerca in SecaaS, in particolare i sistemi SIEM. Inoltre, si discute i requisiti e le caratteristiche di maggiore interesse relative all'implementazione del SIEM come servizio.
Uno dei survey più recenti è sicuramente rappresentato da [22]. Questo documento contiene una revisione sistematica effettuata per descrivere lo stato attuale della tecnologia SIEM e quali

[7] https://solutionsreview.com/security-information-event-management/security-information-event-management-vendor-map/

potrebbero essere i prossimi sviluppi futuri. Ci si concentra quindi su dove i sistemi SIEM si sposteranno nel futuro a breve/lungo termine, se questo cambiamento influenzerà la tecnologia così com'è adesso, ed infine quali vantaggi otterranno gli utenti da questa crescente tecnologia di monitoraggio della sicurezza. Il paradigma di questa tecnologia sta lentamente virando verso esigenti standard internazionali, ai quali tutti gli strumenti di sicurezza devono conformarsi in ogni audit interno o esterno. Gli obiettivi dei SIEM sono rivolti a migliorare i motori di rilevamento per farli rispondere più velocemente e in modo più agile e accurato, ottimizzando così i tempi dell'analista. Nel documento si evidenzia inoltre l'utilizzo insieme a tecnologie all'avanguardia come Blockchain, sistemi per Container, e Cloud. Gli autori utilizzano i documenti analizzati come base su cui proporre un nuovo framework compatibile con il GDPR, utilizzando tecnologie Blockchain, crittografia e Container. Tale framework è stato denominato SIEM-SC (*Security Compliance*). In questo senso, una proposta precedente è rappresentata da [23].
L'articolo in [24] presenta una sistematizzazione dei metodi di correlazione degli eventi di sicurezza in diverse categorie, come l'anno di pubblicazione, i metodi di correlazione applicati, i metodi di estrazione della conoscenza, le fonti di dati utilizzate, le soluzioni architetturali e la valutazione della qualità dei metodi di correlazione. Il lavoro in [25] si concentra invece sulle tecniche di Intelligenza Artificiale per correlazione di eventi, utilizzando modelli basati su rule-based, semantic, graphical e machine learning.

Una lista di sistemi SIEM

Di seguito riportiamo 29 sistemi SIEM, per la maggior parte strumenti commerciali. Dove non specificato, il costo non è direttamente pubblicizzato ma c'è bisogno di entrare in contatto con il reparto vendite per avere un preventivo e più informazioni per l'acquisto.

1. IBM Security QRadar
2. Splunk Enterprise Security
3. LogRhythm NextGen SIEM Platform
4. Sumo Logic
5. AlienVault USM
6. LogPoint
7. Microsoft Azure Sentinel
8. Rapid7 InsightIDR (il costo per InsightIDR Advanced parte da 5,61$/mese per risorsa)
9. McAfee Enterprise Security Manager
10. Datadog (lo strumento offre una prova gratuita dopo la quale si può optare per 0,20$ per GB di log analizzati, al mese. Può anche essere fatturato annualmente o 0,30$ su richiesta)
11. Blumira cloud SIEM platform (Lo strumento offre una vasta gamma di opzioni di prezzo tra cui scegliere, da una prova gratuita a un livello avanzato. Per Microsoft 365, si deve pagare 8$/utente, al mese; Il cloud costa 12$/utente al mese; e l'opzione Avanzata costa 1,6$/utente, al mese)
12. Graylog
13. FortiSIEM
14. Juniper Secure Analytics
15. EventSentry
16. Advanced Security Manager by Cisco
17. Micro Focus ArcSight Enterprise Security Manager (ESM)
18. Panther
19. Prelude
20. SmartEvent event management
21. Next-Gen SIEM from Logsign

22. Securonix Security Operations
23. Devo
24. Exabeam Security Management
25. Snare
26. FireEye Helix
27. BMC AMI Command Center for Security
28. ManageEngine Log360
29. WhiteRock Cybersecurity

Ci concentriamo adesso sui sistemi che implementano o portano ad avere un SIEM open source e di libero utilizzo (chi più chi meno).

- **OpenSearch** è un progetto software open source lanciato nel 2021 come fork dei progetti Elasticsearch e Kibana, con lo sviluppo guidato da Amazon Web Services. Il progetto include un database (chiamato anche OpenSearch) e la visualizzazione e l'analisi del frontend chiamate OpenSearch Dashboards. Nel gennaio 2021, Elastic, la società dietro l'Elastic Stack (costituita dai progetti Elasticsearch, Kibana, Beats e Logstash e spesso nota come ELK Stack o Elastic Stack) ha annunciato che sarebbe passata a una struttura a doppia licenza basata sulla Server Side Public License (SSPL) e sulla Elastic License, nessuna delle quali è stata però riconosciuta come licenza open source dall'Open Source Initiative (OSI). In risposta, Logz.io ha collaborato con Amazon e altri leader del settore per creare un'alternativa open source al nuovo stack ELK closed source, ed è nato OpenSearch.
- **ELK** è probabilmente lo strumento open source più popolare utilizzato come elemento costitutivo in un sistema SIEM. Lo stack ELK è composto da Elasticsearch, Logstash, Kibana e dalla famiglia di spedizionieri di log Beats. Elasticsearch e Kibana sono sotto licenze SSPL a partire dal 14 gennaio 2021. Non è più quindi uno strumento open source in senso stretto, ma è sempre utilizzabile in modo gratuito. Logstash è un aggregatore di log in grado di raccogliere ed elaborare dati da quasi tutte le fonti di dati. Può filtrare, elaborare, correlare e in generale migliorare tutti i dati di registro che raccoglie. Elasticsearch è il motore di archiviazione e una delle migliori soluzioni nel suo campo per l'archiviazione e l'indicizzazione di dati di serie temporali. Kibana è il livello di visualizzazione dello stack. Beats include una varietà di datashipper di log leggeri che sono responsabili della raccolta dei dati e della loro spedizione nello stack tramite Logstash.
- **Logz.io Cloud SIEM** è basato su OpenSearch. Lo strumento offre tre opzioni di prezzo: Comunità che è gratuita. L'opzione Pro è disponibile a 0,92$ e per l'opzione Enterprise è necessario chiedere un preventivo al team di vendita. Open 360TM è la piattaforma di analisi di sicurezza di Logz.io, che unifica l'analisi di log, metriche e tracce. Esso è tutto basato su tecnologie di osservabilità open source come OpenSearch, OpenTelemetry, Prometheus e Jaeger. La piattaforma Logz.io 360 è una piattaforma di osservabilità completa che include tutte le funzionalità della piattaforma Logz.io, tra cui registrazione, metriche, tracciamento e analisi della sicurezza. La piattaforma 360 è progettata per aiutare le organizzazioni a monitorare e risolvere i problemi delle loro applicazioni e infrastrutture, oltre a ottenere informazioni dettagliate sulle loro operazioni. D'altra parte, Logz.io Cloud SIEM è un prodotto specifico all'interno della piattaforma Logz.io che si concentra sull'analisi della sicurezza e sul rilevamento delle minacce. Cloud SIEM raccoglie e analizza i dati di registro da tutta l'infrastruttura di un'organizzazione per identificare potenziali minacce alla sicurezza e fornire avvisi in tempo reale. Cloud SIEM include diverse funzionalità di sicurezza, come il rilevamento delle intrusioni, l'intelligence sulle minacce e la gestione della conformità, per aiutare le organizzazioni a proteggere i propri sistemi e dati. In sintesi, mentre Logz.io 360 Platform e Logz.io Cloud SIEM fanno parte della più ampia piattaforma di osservabilità Logz.io, la piattaforma 360 è una soluzione di

osservabilità completa, mentre Cloud SIEM si concentra specificamente sull'analisi della sicurezza e sul rilevamento delle minacce.

- **OSSIM**: consiste nella versione open source dell'offerta Unified Security Management (USM) di AlienVault. OSSIM è probabilmente una delle piattaforme SIEM open source più popolari, ed include componenti SIEM chiave, vale a dire la raccolta, l'elaborazione e la normalizzazione degli eventi. L'elenco dei progetti open source inclusi in OSSIM include: FProbe, Munin, Nagios, NFSen/NFDump, OpenVAS, OSSEC, PRADS, Snort, Suricata e TCPTrack. L'inclusione di OpenVAS è di particolare interesse, poiché OpenVAS viene utilizzato anche per la valutazione della vulnerabilità correlando i registri IDS con i risultati dello scanner di vulnerabilità. Come ci si aspetterebbe, l'OSSIM open source non è così ricco di funzionalità come la sua alternativa commerciale ed entrambi soffrono di problemi di scalabilità. Le funzionalità di gestione dei log nella versione open source di OSSIM sono praticamente inesistenti.
- **Prelude**: simile a OSSIM, Prelude è un framework SIEM che unifica vari altri strumenti open source. Ancora come OSSIM, esso consiste in una versione open source dell'omonimo strumento commerciale. Prelude accetta registri ed eventi da più fonti e li archivia tutti in un'unica posizione utilizzando l'IDMEF (Intrusion Detection Message Exchange Format). Fornisce funzionalità di filtraggio, rilevamento, avviso, analisi e visualizzazione. Come OSSIM, la versione open source di Prelude è limitata rispetto all'offerta commerciale in tutte queste funzionalità. Citando la documentazione ufficiale: "Prelude OSS è destinato a scopi di valutazione, ricerca e test su ambienti ristretti. Si prega di notare che le prestazioni di Prelude OSS sono molto inferiori rispetto all'edizione Prelude SIEM".
- **Apache Metron**: evolvendosi dalla piattaforma OpenSOC di Cisco e rilasciato per la prima volta nel 2016, Apache Metron è un *data lake* e non uno strumento SIEM open source di per sé. Esso consiste in un altro esempio di framework di sicurezza che combina più progetti open source in un'unica piattaforma. Dal punto di vista dell'architettura, Metron si affida ad altri progetti Apache per la raccolta, lo streaming e l'elaborazione dei dati di sicurezza. Le sonde Apache Nifi e Metron raccolgono dati da fonti di dati di sicurezza che vengono poi inseriti in Apache Kafka. Gli eventi vengono successivamente analizzati e normalizzati in JSON standard e quindi arricchiti e in alcuni casi etichettati. Gli avvisi possono essere attivati se vengono identificati determinati tipi di evento. Per la visualizzazione, le distribuzioni Metron usano comunemente Kibana. Il progetto sembra però essere stato ritirato da Apache a Dicembre 2020.

Fonti OSINT

Con l'acronimo **OSINT**, che sta per **Open Source INTelligence**, ci si riferisce tipicamente al complesso di attività di ricerca, raccolta, catalogazione, analisi e disseminazione di informazioni da *fonti aperte* a scopi di intelligence. A volte il termine OSINT viene usato per indicare le informazioni stesse prodotte da tali attività di intelligence partendo da fonti aperte. È importante non confondere l'OSINT con il software libero, chiamato in inglese *open source software*.

Le origini dell'OSINT risalgono al periodo della seconda guerra mondiale [26], quando alcune agenzie di sicurezza nazionale usufruirono di fonti pubbliche per ottenere informazioni utili alla causa militare. Un noto esempio documentato fu quello dell'agenzia FBMS (Foreign Broadcast Monitoring Service), istituita dal governo statunitense nel 1941: l'FBMS riuscì a correlare le variazioni del prezzo

delle arance a Parigi con i bombardamenti di ponti ferroviari durante la seconda guerra mondiale [27].

Nell'era digitale, con l'avvento di Internet, dei social media e delle grandi piattaforme per la condivisione di contenuti multimediali, l'importanza e l'efficacia dell'OSINT è sensibilmente aumentata. Nella guerra in Ucraina, l'OSINT è stata utilizzata per avere un'idea più precisa dei movimenti dei mezzi militari russi e per comprendere i progressi reali durante la guerra. Tra le fonti utilizzate compaiono filmati di prima mano raccolti da comuni cittadini, così come dati raccolti da siti Web che cercano di tracciare i movimenti di aerei e treni [28, 29].

Va osservato che non esiste una definizione universalmente riconosciuta di OSINT e una sua caratterizzazione come disciplina di intelligence, si tratta infatti di questioni ancora oggetto di dibattito [26]. Il concetto stesso di *fonte aperta* (*open source information*) è stato definito in modi diversi e con variazioni significative nel corso del tempo [30, 31] a causa soprattutto della rapida evoluzione delle tecnologie digitali. In modo molto approssimativo, si può affermare che il termine *fonti aperte* si riferisca ad informazioni non classificate ovvero non coperte da segreto di stato, cosa che può variare da paese a paese in funzione della legislazione vigente. Una definizione meno restrittiva considera come *fonti aperte* tutte quelle la cui raccolta non implichi alcun tipo di tecnica clandestina [32].

Scopo di questo capitolo è fornire una descrizione generale del mondo OSINT, comprensiva di una disamina delle tipologie di fonti, del tipico flusso di lavoro degli analisti OSINT, delle principali tecnologie e strumenti attualmente in uso. Sarà inclusa un'analisi della letteratura scientifica e delle riviste specializzate del settore. Per rendere la trattazione il più possibile auto-contenuta la prima sezione è dedicata a fornire le definizioni e la terminologia comunemente adottate in ambito OSINT. Si noti inoltre che buona parte dei contenuti di questo capitolo, e relativi approfondimenti (sempre redatti in italiano), sono reperibili in un documento tecnico della Facoltà di Scienze e Tecnologie dell'Università degli Studi Camerino, dal titolo "*Il mondo dell'Osint*" [33].

Abbreviazioni e acronimi

CIA	Central Intelligence Agency
FBIS	Foreign Broadcast Information Service
FBMS	Foreign Broadcast Monitoring Service
GL	Gray Literature
IC	Intelligence Community
IDGC	InDividually Generated Content
ISGC	InStitutionally Generated Content
L-FSMC	Long-Form Social Media Content
NMC	News Media Content
NMD	New Media Data
OSD	Open Source Data
OSIF	Open Source InFormation
OSINT	Open Source INTelligence
OSINT-V	Validated Open Source INTelligence
S-FSMC	Short-Form Social Media Content

Definizioni e terminologia

Una delle prime definizioni ufficiali del termine OSINT fu data da una legge statunitense del 2006, specificatamente dalla Public Law 109-163 [34], la quale recita: "*Open-source intelligence (OSINT) is*

intelligence that is produced from publicly available information and is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement.".

Una definizione simile viene fornita dal database ufficiale di terminologia della NATO [35]: "*Intelligence derived from publicly available information, as well as other unclassified information that has limited public distribution or access*.".

Mark M. Lowenthal, ex vice-direttore della Central Intelligence for Analysis, ha fornito una definizione di OSINT che specifica più chiaramente il tipo di fonti utilizzate [31]: "*any and all information that can be derived from overt collection: all types of media, government reports and other documents, scientific research and reports, commercial vendors of information, the Internet, and so on. The main qualifiers to open-source information are that it does not require any type of clandestine collection techniques to obtain it and that it must be obtained through means that entirely meet the copyright and commercial requirements of the vendors where applicable.*".

Su questa falsariga, il politologo americano Jeffrey T. Richelson propose nel 2016 una definizione tesa a chiarire ulteriormente il concetto di *fonti aperte* [36]: "*open source acquisition involves procuring verbal, written, or electronically transmitted material that can be obtained legally. In addition to documents and videos available via the Internet or provided by a human source, others are obtained after U.S. or allied forces have taken control of a facility or site formerly operated by a foreign government or terrorist group.*".

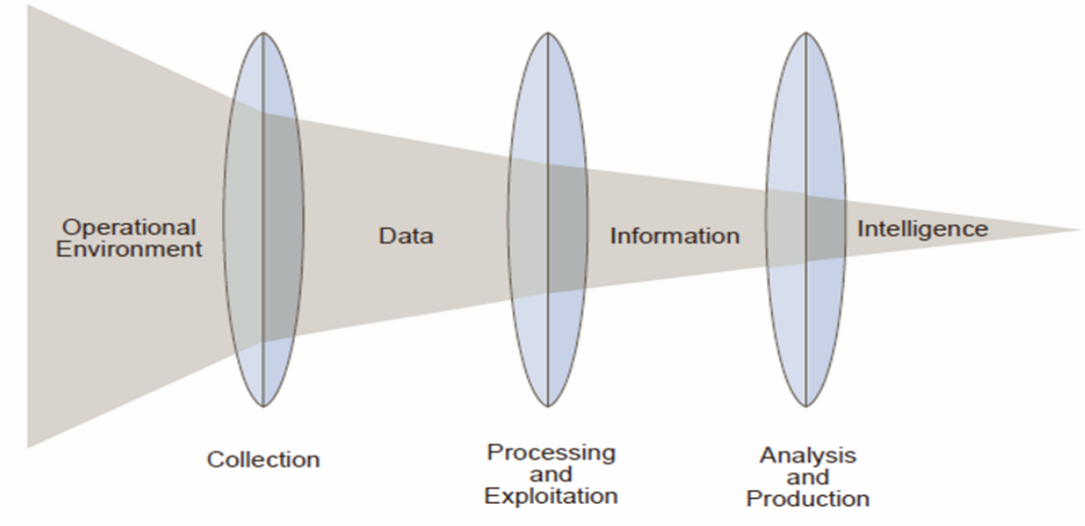

Figura 3: Relazione tra dati, informazioni e intelligence [37]

Alcuni analisti [26] sostengono che i grandi cambiamenti tecnologici dell'ultimo ventennio impongono una distinzione tra *OSINT di prima generazione* e *OSINT di seconda generazione*, il cui inizio è collocabile intorno al 2005, grossomodo coincidente con l'avvento del Web 2.0. Viene fornita pertanto una definizione di OSINT di seconda generazione, coerente con la Public Law 109-163, che recita: "*We define OSINT as publicly available information that has been discovered, determined to be of intelligence value, and disseminated by a member of the intelligence community (IC)*".

Nel presente documento si farà riferimento alla seconda generazione di OSINT.

Sempre in [26] viene fatta una distinzione netta tra **Open Source InFormation** (**OSIF**) e OSINT, per OSIF si intendono generici dati non classificati e pubblicamente disponibili, mentre l'OSINT è il risultato prodotto da uno sforzo di elaborazione e di sfruttamento delle informazioni al fine di convalidarne la rilevanza, l'accuratezza e l'utilizzabilità.

Si sottolinea comunque che non esiste una definizione universale di OSINT e OSIF. L'autorevole libro di testo *Handbook of Intelligence Studies* [38] identifica quattro distinte categorie di dati in ambito OSINT, relative alle varie fasi in cui si articola l'attività di intelligence:

- **Open Source Data** (**OSD**): stampa grezza, trasmissioni, resoconti orali o altre forme di informazioni provenienti da una fonte primaria. Possono essere fotografie, registrazioni su nastro, immagini satellitari, lettere personali, post online, ecc.
- **Open Source Information** (**OSIF**): informazioni generiche tipicamente destinate a un'ampia diffusione che combinano più dati utilizzando un certo livello di validazione. Ad esempio, libri, giornali e notiziari.
- **Open Source Intelligence** (**OSINT**): informazioni che sono state deliberate, scoperte, discriminate, distillate e diffuse a un pubblico selezionato.
- **Validated Open Source Intelligence** (**OSINT-V**): informazioni caratterizzate da un elevato grado di attendibilità. Quest'ultime possono a loro volta suddividersi in due sotto-categorie:
    - informazioni che provengono da una fonte consolidata e affidabile e/o che possono essere convalidate da un confronto con altri dati;
    - informazioni che possono essere ritenute valide nel loro formato originale, come ad esempio i notiziari che mostrano il discorso di un leader statale. Ovviamente, va considerata la possibilità di manipolazione o falsificazione.

Figura 4: Ciclo di trasformazione dei dati nell'OSINT, dai dati grezzi di fonti aperte all'informazione di intelligence validata [39]

Una classificazione alternativa delle fonti aperte aderente alla seconda generazione di OSINT è fornita in [26], dove viene fatta preliminarmente una distinzione tra OSD e OSIF:

- **Open Source Data** (**OSD**): si tratta di tutti quei dati aperti che individualmente hanno uno scarso valore, ma che in forma aggregata e contestualizzata potrebbero avere un valore significativo per l'intelligence. Si pensi ad esempio ad un singolo tweet su Twitter inneggiante al terrorismo islamico. Di per sé tale dato è irrilevante per l'intelligence, ma un'intera collezione di tweet simili provenienti da un'area geografica circoscritta potrebbe costituire un'informazione di grande rilievo per l'intelligence. Va rimarcato che gli OSD includono tutto quel materiale pubblico che non è stato ancora esplicitamente diramato, ma che è ancora pubblicamente e commercialmente disponibile, come ad esempio le immagini satellitari commerciali.
- **Open Source InFormation** (**OSIF**): materiale ottenibile legalmente tramite richiesta, acquisto o l'osservazione di un normale cittadino. Oltre agli OSD sono pertanto inclusi anche altri contenuti di potenziale interesse per l'intelligence. L'OSIF è quindi la categoria più ampia di informazioni disponibili pubblicamente o commercialmente.

Nella seguente sezione sarà offerta una disamina più accurata delle tipologie di fonti OSINT.

Tipologie di fonti OSINT

Richelson [36] individua sei diverse categorie di flussi informativi utilizzabili in ambito OSINT:

- **Media**: giornali stampati, riviste, radio e televisione da un paese all'altro e tra paesi diversi.
- **Internet**: pubblicazioni online, blog, gruppi di discussione, citizen media (ad esempio, video dei cellulari e contenuti creati dagli utenti), YouTube e altri siti web di social media (ad esempio, Facebook, Twitter, Instagram, ecc.). Questa fonte supera anche una serie di altre fonti in termini di tempestività e facilità di accesso.
- **Public government data**: rapporti governativi pubblici, bilanci, audizioni, elenchi telefonici, conferenze stampa, siti web e discorsi. Sebbene queste informazioni provengano da una fonte ufficiale, sono pubblicamente accessibili e possono essere utilizzate apertamente e liberamente.
- **Professional and academic publications**: informazioni acquisite da riviste specialistiche e scientifiche, da conferenze, simposi, articoli accademici, tesi e tesine.
- **Dati commerciali**: immagini e database commerciali, valutazioni finanziarie e industriali.
- **Grey literature**: relazioni tecniche, preprint, brevetti, documenti di lavoro, documenti aziendali, lavori non pubblicati e newsletter.

Sempre Richelson osserva che l'OSINT va distinta dalla semplice ricerca di informazioni da fonti aperte, in quanto viene applicato un rigoroso processo di intelligence per selezionare le fonti e per verificarne il grado di attendibilità. L'obiettivo è colmare il gap informativo per supportare scelte decisionali di tipo strategico di singoli e/o di gruppi.

In [26] viene proposta una diversa classificazione delle fonti OSIF che risulta più attinente alla seconda generazione di OSINT e in particolare alle relative fasi di indagine – il cosiddetto ciclo OSINT– che saranno illustrate nel seguito. Tale classificazione divide le tipologie di fonti aperte in due categorie di primo livello, ciascuna delle quali è a sua volta suddivisa in due categorie di secondo livello.

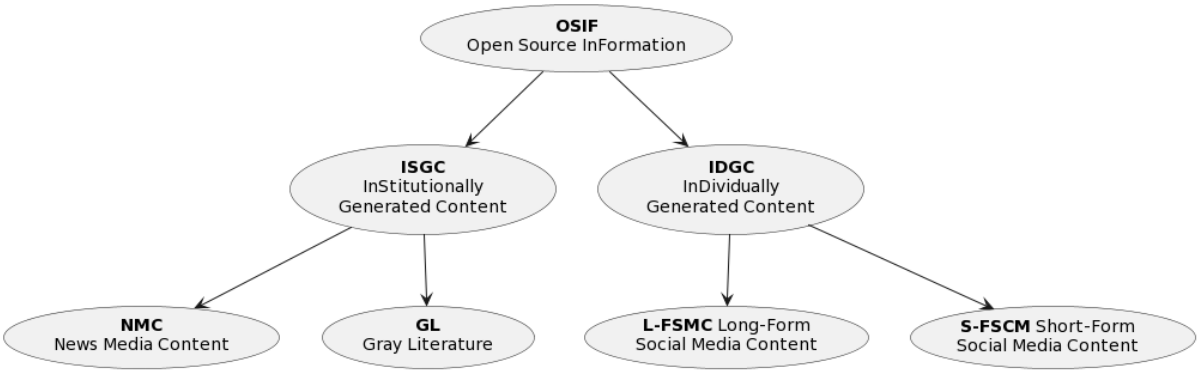

Figura 5: Classificazione delle fonti aperte (OSIF) fornita in [26]

La prima distinzione è fatta in base al soggetto/organizzazione che genera i contenuti: contenuti generati da un qualche tipo di istituzione (**InStitutionally Generated Content ISGC**) o contenuti generati da singoli individui (**InDividually Generated Content IDGC)**. I contenuti generati dalle istituzioni sono prodotti dai mezzi di informazione o da altre organizzazioni riconosciute, molti dei quali possono essere stati definiti in precedenza come letteratura grigia. I contenuti generati individualmente, o i contenuti dei social media, si dividono in long-form e short-form, e presentano differenze importanti per l'elaborazione e l'utilizzo.

Segue una breve descrizione delle categorie di secondo livello.

- **News Media Content NMC** Il contenuto dei mezzi di informazione è auto-identificato e pubblicamente riconosciuto come giornalismo. Le sue fonti sono multimediali: riviste (sia cartacee che online), televisione e radio. I mezzi di informazione comprendono anche i siti di aggregazione di notizie, che possono pubblicare o meno contenuti originali. Sono inclusi in questa categoria anche i contenuti prodotti dallo Stato a patto che vengano diramati da un organo di informazione.
- **Gray Literature GL** Per letteratura grigia si intendono quei contenuti provenienti da organizzazioni ed istituzioni non mediatiche, sia pubbliche che private. Viene incluso in questa categoria il materiale proveniente da università, centri di ricerca, editori privati, governi nazionali, gruppi di esperti, associazioni di categoria e sindacati. Si presume che una parte rilevante dei contenuti istituzionali non sia presente nel cyberspace, ma che sia ancora gestita in modo manuale e non digitalizzato.
- **Long-Form Social Media Content L-FSMC** Contenuti testuali di elevata lunghezza prodotti da singoli individui o da piccoli gruppi. A titolo di esempio, si pensi ai contenuti di un blog o a quelli di siti Web quali Tumblr e Reddit. Si osservi che gran parte dell'analisi dei contenuti di tipo long-form è trascurata a favore di quelli in forma breve (short-form), tipica dei social media.

- **Short-Form Social Media Content S-FSMC** Materiale di lunghezza limitata ottenibile da social media quali Facebook, Twitter, LinkedIn, ecc., diversamente dai contenuti di tipo long-form, quelli in forma breve hanno tipicamente uno scarso valore di intelligence se considerati individualmente, mentre possono acquistare un valore significativo se vengono opportunamente aggregati. Eccezione fatta per i contenuti in forma breve relativi a specifici account di grande interesse, ad esempio account di persone famose, leader di pensiero, giornalisti di spicco e personalità di governo.

Metodologia e ciclo delle operazioni OSINT

Le informazioni utili ad un analista OSINT costituiscono solo una piccola parte di quelle diffuse e condivise quotidianamente. Discriminare ciò che è più importante ed utile, da ciò che lo è di meno, richiede uno sforzo enorme, distribuito in tutte le fasi in cui si articola l'indagine, dalla raccolta iniziale alla diffusione dei risultati ai responsabili politici che li ricevono. In particolare, il processo di trasformazione di OSD in informazioni utilizzabili richiede una fase cruciale di contestualizzazione dei dati, necessaria per valutarne la validità e l'affidabilità.

Va osservato che non esiste ancora una metodologia chiara e condivisa sul come condurre un'indagine OSINT. Tuttavia, nel corso del tempo sono stati proposti alcuni modelli che tentano di descriverne le fasi salienti. Quasi tutti si ispirano al modello iniziale proposto dalla CIA [40] che va sotto il nome di *ciclo dell'intelligence* (*intelligence cycle*), poiché le varie fasi individuate si susseguono in modo ciclico; si osservi che tale modello si applica a tutte le fonti, non solo a quelle aperte dell'OSINT.

Per la CIA il *ciclo dell'intelligence* (o *processo dell'intelligence*) si compone di cinque fasi:

1. *Pianificazione e Direzione* (*Planning and Direction*): l'attività di intelligence inizia sempre con una fase preparatoria in cui si pianifica cosa deve essere fatto e come. Viene stabilita una direzione specifica dell'indagine entro la quale devono muoversi gli analisti. In funzione delle informazioni in loro possesso e di ciò che si desidera scoprire, si discutono i possibili modi per raccogliere dati e informazioni potenzialmente utili.
2. *Raccolta* (*Collection*): si raccolgono informazioni grezze multimediali (testo, immagini, video, documenti cartacei, ecc.) potenzialmente utili all'indagine, sia da fonti segrete che da fonti aperte.
3. *Elaborazione* (*Processing*): tutte le informazioni grezze raccolte vengono inserite all'interno di un rapporto di intelligence. Tale rapporto potrebbe includere contenuti di natura molto diversa, dalla traduzione di un documento segreto alla descrizione di una foto satellitare.
4. *Analisi e Produzione* (*Analysis and Production*): in questa fase vengono esaminate in modo più dettagliato le informazioni raccolte e organizzate nel rapporto, analizzando in particolare come si combinano, mantenendo il focus sull'obiettivo dell'indagine. Viene valutato attentamente cosa sta accadendo, perché sta accadendo, cosa potrebbe accadere dopo, e come potrebbe influire sugli interessi degli Stati Uniti.

5. *Disseminazione* (*Dissemination*): in questa fase finale, viene redatto un documento con il risultato dell'indagine. Tale documento viene trasmesso ad un decisore politico, lo stesso che ha richiesto l'indagine. Dopo aver letto l'analisi finale e avere appreso la risposta alla domanda iniziale, il decisore politico potrebbe porre ulteriori domande. Quindi l'intero processo ricomincia da capo.

Figura 6: Ciclo delle operazioni di intelligence [37]

In ambito prettamente OSINT, il modello proposto dalla CIA è stato adottato da Gibson [39, 40] e, con alcune correzioni, da Hassan e Hijazi [41]. Mentre un'interpretazione pratica inserita in un manuale di formazione adottato da diverse agenzie governative statunitensi viene data da Bazzell [42, 43]. In [26] viene illustrato un modello concepito per la seconda generazione di OSINT. Tale modello, come mostrato nella Figura 7, prevede quattro passaggi fondamentali:

1. *Raccolta* (*Collection*)
2. *Elaborazione* (Processing)
3. *Sfruttamento* (Exploitation)
4. *Produzione* (*Production*).

In modo approssimativo, tali fasi consistono nell'acquisizione delle informazioni, nella convalida, nell'identificazione dell'utilità e del valore delle informazioni e nella loro fornitura ai committenti.

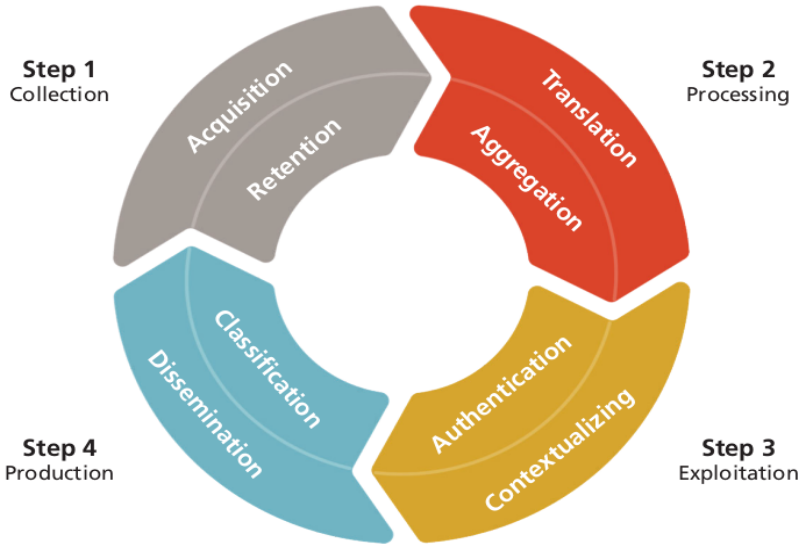

Figura 7: Ciclo delle operazioni nei processi OSINT di seconda generazione [26]

Di seguito forniremo una descrizione più dettagliata di ciascuna fase, scomponendola in sotto attività esattamente come fatto in [26], in quanto reputiamo tale metodologia quella più in linea con le finalità del progetto in oggetto. Sarà inclusa anche una stima del livello di difficoltà di ciascuna sotto-attività del suddetto ciclo metodologico e per ciascuna delle quattro tipologie di fonti OSINT; sarà utilizzata una scala qualitativa a tre valori: *easy*, *medium* e *hard*.

Step 1: Raccolta (Collection)

La prima fase, la *raccolta*, si compone di due sotto-attività: l'*acquisizione* (*acquisition*) e la *conservazione* (*retention*) di informazioni potenzialmente utili. Chiaramente è necessario procedere secondo una direzione coerente con l'obiettivo dell'indagine, sia per quanto concerne la scelta dei tipi di informazioni da raccogliere, sia per quanto concerne l'assegnazione di un adeguato livello di priorità. Più specificatamente, per *acquisizione* (*acquisition*) si intendono tutte quelle operazioni che è necessario espletare per ottenere determinate informazioni, sia in formato fisico che in formato elettronico. La *conservazione* (*retention*) consiste invece nell'insieme di attività necessarie al corretto mantenimento delle informazioni acquisite.

Con riferimento ai quattro tipi di OSIF relativi alla seconda generazione di OSINT, i news media content (NMC) sono i più semplici da raccogliere. L'acquisizione richiede infatti uno sforzo minimo, trattandosi nella maggioranza dei casi di contenuti disponibili on-line. Similmente, la conservazione non nasconde particolari insidie, poiché il volume di tali informazioni è gestibile e i formati utilizzati sono spesso di tipo testuale e standard. Pure la raccolta di gray literature (GL) si sta semplificando molto negli ultimi anni, per ragioni analoghe. Tuttavia, siccome i creatori di letteratura grigia si stanno digitalizzando più lentamente, non di rado si è costretti a raccogliere informazioni in formato cartaceo, soprattutto quando si ha a che fare con paesi in via di sviluppo, caratterizzati da una scarsa diffusione della rete Internet. Analogamente ai news media content, anche nel caso della gray literature la conservazione non presenta particolari criticità.

La raccolta risulta invece particolarmente sfidante nel caso di social media content (SMC), sia per contenuti di tipo short-form (S-FSMC) che long-form (S-FSMC). In prima istanza, può essere difficile avere un quadro completo dei dati grezzi da raccogliere. Se in passato l'analisi dei social media poteva essere fatta a costi ridotti o addirittura gratuiti, si pensi ad esempio alla piattaforma Topsy, che offriva un accesso pubblico a grandi moli di dati di Twitter, oggi lo scenario è profondamente mutato. L'analisi dei social media costituisce infatti un settore industriale consolidato, e differentemente dal passato è in mano a grandi piattaforme private i cui dati sono chiusi o molto costosi.

Inoltre, le società che aggregano e commercializzano dati di piattaforme di social media forniscono spesso solo una frazione dei dati esistenti in tali piattaforme, e/o dati che si riferiscono ad una specifica finestra temporale. Per di più, tali società si focalizzano tipicamente su piattaforme statunitensi, come Facebook e Twitter, tralasciando piattaforme potenzialmente più interessanti per l'intelligence. Va osservato inoltre che, qualora si acquisisca un set completo di dati dai social media, i dati ottenuti potrebbero non costituire un campione rappresentativo di una popolazione. L'uso dei social media infatti non è distribuito in modo omogeneo nella popolazione, e in alcuni luoghi di interesse per l'intelligence, talune categorie sociali non sono affatto rappresentate.

La conservazione di social media data pone inoltre problemi legali non banali, in funzione della legislazione vigente sulla protezione dei dati personali nel paese ove sono mantenuti. Si pensi ad esempio alle restrizioni imposte dal GDPR in Europa. Tali problematiche legali sono meno presenti nel caso di gray literature e quasi inesistenti nel caso di news media content.

Altro elemento critico è costituito dalla dinamicità dei contenuti dei social media. Un articolo di cronaca è piuttosto statico nel tempo (salvo eventuali correzioni); un nuovo articolo viene creato se vi sono cambiamenti sostanziali in una storia. Invece, un argomento di tendenza in un forum online può suscitare interesse e quindi aggiornamenti per giorni, settimane, o addirittura anni. Nel caso specifico di social media content, l'acquisizione e la conservazione devono avvenire in modo continuativo e in tempo reale, perché contenuti di rilevante interesse per l'intelligence potrebbero essere postati e subito rimossi se dovessero causare polemiche, controversie o rivelare informazioni sensibili.

In ultimo, va osservato che c'è un crescente aumento di contenuti non testuali sia di tipo L-FSMC che S-FSMC. Si pensi ad esempio ai video caricati in YouTube, che rappresentano L-FSMC non testuali, oppure alle immagini caricate nella piattaforma Flickr o ai video "live" di piattaforme quali Facebook e Twitter, tutti esempi quest'ultimi di S-FSMC non testuali. Tali contenuti sono chiaramente di più difficile gestione rispetto a quelli di tipo puramente testuale.

Step 2: Elaborazione (Processing)

L'*elaborazione* (*processing*) costituisce la seconda fase del ciclo OSINT ed è finalizzata a validare le informazioni e a renderle disponibili. Viene attuata eseguendo varie attività, quali la traduzione di contenuti dalla lingua di origine all'inglese, e la trasformazione di materiale video fotografico in intelligence utilizzabile.

Ovviamente, con l'avvento dell'OSINT di seconda generazione, la fase di elaborazione ha subito un cambiamento epocale rispetto a come veniva condotta nella prima generazione di OSINT. Sono state infatti introdotte modifiche sostanziali ai metodi preesistenti, e nuovi metodi sono sorti. Molte attività oggi possono essere eseguite in modo più rapido e a costi ridotti grazie ai software di ultima generazione per il processamento dei linguaggi naturali, si pensi ad esempio a Google Translate, a

DeepL e soprattutto a ChatGPT [44]. Al tempo stesso, vi è un'abbondanza di informazioni disponibili in formati destrutturati o debolmente strutturati, cosa che inevitabilmente aumenta la complessità. L'elaborazione consiste prevalentemente di due sotto attività: la *traduzione* (*translation*) e l'*aggregazione* (*aggregation*). Tali attività possono anche avvenire in parallelo, sebbene in certi casi una può essere di supporto all'altra.

Come già detto, l'elaborazione di news media content (NMC) in lingua straniera richiede la traduzione in inglese. Se un tempo ciò costituiva il grosso dello sforzo del *Foreign Broadcast Information Service* (*FBIS*), oggi la traduzione ha enormemente beneficiato dei rapidi progressi dei sistemi di traduzione automatici; almeno per le lingue di cui si dispone di sintassi documentate e di corpus di documenti per l'addestramento di modelli di machine learning. Sebbene i linguisti professionisti giochino ancora un ruolo centrale, aggiungendo sfumature e un contesto culturale ai contenuti in lingua straniera, possono ora spostare i loro sforzi verso la fase di sfruttamento (exploitation), fornendo un valore analitico alle informazioni ottenute.

I sistemi di traduzione automatica offrono migliori prestazioni per news media content (NMC), poiché utilizzano un vocabolario standard e spesso hanno una struttura regolare e ripetitiva. La gray literature (GL) tipicamente segue standard di scrittura professionale che favorirebbero la traduzione automatica, tuttavia, taluni argomenti avanzati e molto specifici richiedono spesso l'intervento umano.

Ci sono invece sia vantaggi che svantaggi per quanto concerne i contenuti informativi dei social media. Da un lato, la lunghezza dei post è tipicamente limitata. Nel caso di Twitter, vi è stato per molto tempo un limite massimo di 140 caratteri, al momento il limite è di 280 caratteri per gli utenti non registrati, e di 4000 caratteri per gli utenti registrati negli USA. Ci sono invece sia vantaggi che svantaggi per quanto concerne i contenuti informativi dei social media. Da un lato, la lunghezza dei post è tipicamente piuttosto limitata. Dall'altro lato, i post nei social media non di rado contengono espressioni gergali, forme di stenografia (per nascondere il reale contenuto di un messaggio), emoji o icone. I post possono inoltre mischiare più linguaggi e frequentemente contengono errori tipografici. Mentre i long-form social media content possono contenere sufficienti informazioni ad ottenere una registrazione coerente per desumere la posizione e lo stile dell'autore, è molto più improbabile che i contenuti di tipo short-form forniscano una siffatta registrazione, a meno che non venga compilato un vero e proprio corpus di materiale.

L'attività di aggregazione, tipicamente non necessaria per contenuti di tipo GL (gray literature) e NMC (news media content), costituisce invece una fase critica per l'analisi di molti tipi di MSC (social media content), in particolare quelli in forma breve (short-form). L'aggregazione può anche includere eventuali riduzioni e/o integrazioni quando si traduce un corpo di dati in una forma utilizzabile. Va osservato inoltre che vi sono molte società commerciali che offrono servizi di aggregazione di dati di social media, cosa che elimina la necessità dell'IC di effettuare una raccolta diretta.

Sebbene tali servizi di aggregazione riducano sensibilmente lo sforzo per la raccolta e per l'elaborazione delle informazioni, potrebbero però non fornire dati di alcune piattaforme, oppure non fornire campioni completi di dati. Un compito non banale inoltre per l'IC è capire esattamente quali contenuti sono stati inseriti nel set di dati fornito dagli aggregatori, cosa che rende molto difficoltosa la successiva contestualizzazione e autenticazione (certificazione della provenienza).

Step 3: Sfruttamento o analisi (Exploitation or analysis)

Lo *sfruttamento*, detto anche *analisi*, serve a determinare il valore che una determinata informazione può avere per l'intelligence. Arthur Hulnik, un esperto di intelligence ed ex ufficiale della CIA, sostiene che uno degli aspetti più sfidanti riguardante l'OSINT è la grande vastità di informazioni pubblicamente disponibili, e la grande variabilità di attendibilità delle stesse. Pertanto, una buona parte del tempo degli analisti OSINT serve a separare le informazioni affidabili, "good intelligence", da quelle non affidabili, "bad intelligence". Secondo Libor Benes [45] gli analisti devono essere in grado di svolgere i seguenti compiti: "*gather, judge, and sort information, know and handle limitations, and understand different users, needs, tasks, information mix, organization, institutions, and the law*". Il risultato dell'analisi dovrebbe includere conclusioni analitiche guidate dalle fonti disponibili. La fase di sfruttamento è generalmente suddivisa in tre sotto fasi: *autenticazione*, *valutazione della credibilità* e *contestualizzazione*.

Per *autenticazion*e si intende la verifica che una determinata informazione provenga realmente da una determinata fonte; fonte che può essere dichiarata esplicitamente, quindi associata all'informazione stessa, oppure presunta in modo più o meno implicito. Nel caso di informazioni provenienti da fonti istituzionali, accertarne la provenienza è un compito abbastanza semplice. È molto improbabile infatti che un articolo sul New York Times venga pubblicato all'insaputa del direttore o di componenti dello staff di tale quotidiano, o che includa contenuti non approvati da quest'ultimi. Similmente, la gray literature pubblicamente disponibile presso i siti Web governativi può essere considerata, con un elevato livello di confidenza, come prodotta e diffusa dal governo. Autenticare i contenuti dei social media è invece molto più difficile. L'identità degli autori dei contenuti può essere deliberatamente nascosta o addirittura falsificata. Ciò non riguarda soltanto il vero nome degli autori. Ad esempio, un utente potrebbe mentire riguardo alle sue abitudini personali, alla sua posizione, alla sua età, ecc. Per accertare le condizioni meteorologiche di un dato luogo tramite analisi di SMC, è molto importante che gli utenti dei post, inerenti al tema meteorologico, si trovino esattamente in quel luogo. In alcune circostanze, l'autenticazione potrebbe dover svolgersi contestualmente alla fase di aggregazione dei dati per assicurare che un campione o un gruppo di dati non sia erroneamente distorto.

Similmente all'autenticazione, non è difficile *valutare la credibilità* dei mezzi d'informazione tradizionali e della gray literature, ma è estremamente complicato farlo per i social media content. Un indicatore di credibilità tenta di stimare l'affidabilità di un'informazione, esprimibile in termini di trasparenza, assenza di mistificazioni, e verificabilità del collegamento con la sua fonte. Ad esempio, il New York Times pubblica materiale cercando di garantire che il contenuto sia accurato e trasparente riguardo alle sue fonti. Ciò potrebbe essere meno vero per le fonti dei media stranieri, quali i media statali, che potrebbero promuovere contenuti propagandistici. Ciò però non esclude che non si possano comunque ottenere indicazioni utili sulla credibilità delle informazioni che diffondono. Differentemente dagli organi di informazione tradizionali, tipicamente nei social media non si trovano informazioni mediate da terzi. In genere il contenuto proviene dalla fonte stessa. Tuttavia, non sempre le cose stanno in questi termini, e spesso l'originalità della fonte può essere dubbia. Bot, repost e retweet possono offuscare e distorcere in modo significativo le intenzioni della fonte originale. Rimangono inoltre tutti gli elementi di soggettività che caratterizzano un'informazione prodotta da un singolo senza alcun tipo di mediazione giornalistica.

Tramite la *contestualizzazione* l'analista di fonti aperte può sfruttare la sua competenza in materia per arricchire e migliorare la qualità dei contenuti. Ciò può consistere nell'aggiunta di commenti riguardanti la fonte e di altri dati che aumentino la credibilità dei contenuti. La contestualizzazione

può anche comportare l'unione di più elementi di OSIF di diversa provenienza in un unico prodotto che fornisca una visione d'insieme e più comprensibile di una data questione.

Step 4: Produzione (Production)

La fase finale di *produzione* (*production*) serve a fornire le informazioni al cosiddetto *consumatore* in una forma utilizzabile. Il *consumatore* sarà tipicamente un analista di intelligence di tipo "all-source", cioè un analista che si occupa di incorporare il risultato in una produzione di tipo "multi-intelligence" che si basa su tutte le fonti, non soltanto quelle aperte. Tuttavia, in alcuni casi il risultato di un processo OSINT può anche avere elevata priorità ed essere sufficientemente completo da essere fornito direttamente ad un decisore politico o ad altri clienti dell'intelligence. In fase di produzione viene anche assegnato un livello di *classificazione* al prodotto OSINT. Sebbene quest'ultimo sia stato ottenuto a partire da fonti aperte (OSIF), eventuali dettagli sulle fasi di raccolta, elaborazione, e sfruttamento delle informazioni possono giustificare un aumento del livello di classificazione. Nel combinare l'OSIF con altre informazioni possono applicarsi specifici requisiti che impongono una forma di riservatezza. Ad esempio, potrebbe essere compromessa l'accessibilità futura a tutte quelle fonti che sono dichiarate nel prodotto OSINT. Inoltre, i produttori di OSIF potrebbero utilizzare tecnologie di sfruttamento ed elaborazione classificate in modo da giustificare la classificazione delle informazioni.

La *disseminazione* (*dissemination*) è un'altra attività che viene svolta in fase di produzione. L'analisi open-source viene tipicamente disseminata sotto forma di rapporti scritti. Non sono escluse però forme di disseminazioni alternative quali briefing verbali o visualizzazioni grafiche. Sfortunatamente, il mezzo di disseminazione più spesso utilizzato è quello più semplice, anziché quello più efficace. Video, audio e visualizzazioni grafiche interattive potrebbero essere molto più efficaci di rapporti scritti nel comunicare determinate informazioni. Gli analisti di intelligence di tipo all-source, ottengono generalmente i rapporti da database testuali, quali Trident, WISE, o Pathfinder. Similmente, i clienti dell'intelligence ricevono spesso i risultati di un'indagine sotto forma di libro di briefing stampato. Va osservato però che i recenti avanzamenti tecnologici dei portali open-source e la transizione dal Presidential Daily Brief al formato iPad stanno aprendo la strada a meccanismi più efficaci nel comunicare informazioni, quali strumenti visuali e file dinamici.

Principali tecnologie utilizzate dall'OSINT

Di seguito riportiamo una lista di alcune tecnologie utilizzate in ambito OSINT e frequentemente menzionate in siti Web specialistici (l'ordine utilizzato è puramente casuale). In alcuni casi si tratta più propriamente di tecnologie classificabili come *Cyber Threat Intelligence tools* (*CTI tools*) includenti alcune funzionalità OSINT. Strumenti di questo tipo sono molto utilizzati in ambito IT, in particolare in cybersecurity, per individuare informazioni di interesse sul sistema informatico di una determinata organizzazione, quali le porte e i servizi esposti, eventuali errori di configurazione, la mappatura della rete, ecc.

Per ciascuno strumento sarà data successivamente una breve descrizione.

- Maltego
- Shodan
- TheHarvester
- Recon-ng

- Spiderfoot
- OSINT Framework
- Foca
- Metagoofil
- GHunt
- Yandex Images
- N2YO.com
- TinEye
- Have I Been Pwned

Maltego è uno strumento OSINT potente e sofisticato per la raccolta di dati da fonti pubbliche. Sviluppato da Paterva, Maltego OSINT consente agli utenti di scoprire rapidamente le relazioni tra grandi quantità di dati di varia natura, utilizzabili per creare profili di intelligence attendibili. Con Maltego OSINT, gli utenti sono in grado di estrarre informazioni da più fonti online utilizzando semplici rappresentazioni grafiche. Ciò include la capacità di mappare i social network, acquisire dettagli di contatto e dati aziendali, tenere traccia di nomi di dominio e indirizzi IP, scoprire prove digitali come documenti o immagini archiviate su siti Web, trovare articoli di notizie correlati e altro ancora. Inoltre, essendo in grado di automatizzare il processo di raccolta dei dati pubblicamente disponibili, Maltego OSINT consente agli utenti di scoprire rapidamente connessioni nascoste che altrimenti rimarrebbero non rilevate.

Shodan è un motore di ricerca per dispositivi Internet of Things (IoT) e uno strumento OSINT utilizzato per scoprire dispositivi vulnerabili ed esposti connessi a Internet. Shodan è stato creato da John Matherly nel 2009 ed è considerato il primo motore di ricerca per computer al mondo. Shodan può essere utilizzato per rilevare vulnerabilità di sicurezza su siti Web pubblici, oltre a fornire informazioni dettagliate su ogni server Web che trova. Shodan è diventato sempre più popolare tra i professionisti della sicurezza informatica e della sicurezza IT che lo utilizzano per la valutazione della vulnerabilità, i test di penetrazione e la mappatura della rete. Shodan li aiuta anche a identificare servizi non sicuri come database cloud configurati in modo errato, server FTP, server telnet e server SSH che sono esposti su Internet senza autenticazione o crittografia. Inoltre, Shodan fornisce informazioni tecniche dettagliate su ogni dispositivo che trova, inclusi indirizzo IP, tipo di sistema operativo, porte aperte, programmi software in esecuzione e vulnerabilità associate.

TheHarvester è un potente strumento OSINT utilizzato per trovare informazioni relative a domini e indirizzi e-mail. Può essere utilizzato da professionisti della sicurezza, amministratori IT e hacker per raccogliere informazioni da diverse fonti su Internet. TheHarvester è stato creato come alternativa per fare ricerca su risorse pubbliche come motori di ricerca, server di chiavi PGP e social network. Consente agli utenti di raccogliere rapidamente grandi quantità di dati da siti come Google, Bing, Yahoo!, Dogpile, LinkedIn, Twitter e molti altri. Tutti i dati raccolti possono essere esportati in diversi formati come HTML/XML o anche salvati come file di testo. Inoltre, include un'API che consente agli utenti di personalizzare le proprie ricerche in base alle proprie esigenze specifiche.

Recon-ng è uno strumento OSINT utilizzato per la ricognizione e la raccolta di dati. È un'applicazione Web completa che può essere utilizzata per raccogliere sottoinsiemi di informazioni pubbliche relative a un target, come nomi utente, nomi, indirizzi e-mail, nomi di dominio e altri dettagli rilevanti. Recon-ng è stato progettato per automatizzare il processo di raccolta di informazioni su un determinato obiettivo nel modo più rapido ed efficiente possibile. Lo strumento Recon-ng OSINT

fornisce agli utenti l'accesso a più risorse come Google, Bing, Twitter, Shodan e altro ancora. La piattaforma consente inoltre agli utenti di interagire con ciascuna risorsa utilizzando la stessa interfaccia che semplifica notevolmente il processo di raccolta dei dati rispetto ai metodi tradizionali. Consente agli utenti di raccogliere rapidamente informazioni complete su un obiettivo senza dover cercare manualmente più fonti o database online.

Spiderfoot è un eccellente strumento OSINT progettato per automatizzare il processo di raccolta informazioni su un obiettivo specifico. Consente agli utenti di accedere in modo rapido e semplice a un'ampia gamma di fonti di dati. È in grado di raccogliere informazioni da oltre 200 fonti, come record DNS, informazioni WHOIS e risorse pubbliche come Shodan, VirusTotal, Google e altri. Spiderfoot può essere utilizzato per ricognizioni, ricerche investigative e persino caccia alle minacce, consentendo agli utenti di identificare rapidamente potenziali minacce o vulnerabilità nel loro ambiente. Lo strumento funziona scansionando Internet alla ricerca di dati disponibili pubblicamente da varie fonti in base ai parametri di query di input dell'utente. I dati raccolti possono quindi essere mappati in un grafico interattivo con vari indicatori visivi che facilitano l'interpretazione delle informazioni raccolte. Questa funzione rende molto più facile per i professionisti della sicurezza riconoscere tendenze o anomalie all'interno delle loro reti che possono aiutarli a rilevare attività o minacce dannose in anticipo.

OSINT Framework è un sito Web e uno strumento di raccolta di informazioni utilizzato dai professionisti della sicurezza per scopi investigativi. È una raccolta di strumenti gratuiti e pubblicamente disponibili che possono essere utilizzati per condurre indagini online. OSINT Framework fornisce agli utenti una piattaforma di facile utilizzo per cercare, raccogliere e analizzare rapidamente dati da varie fonti come piattaforme di social media, siti Web, forum, blog e altro ancora. Utilizzando questo framework, i professionisti della sicurezza sono in grado di raccogliere una grande quantità di informazioni al fine di identificare potenziali minacce o anomalie sul web. OSINT Framework consente agli utenti di accedere ai registri pubblici e ad altre fonti di informazioni in modo rapido ed efficiente. Utilizza motori di ricerca e database specializzati come Google Hacking Database (GHDB) e molti altri strumenti di intelligence open source come Recon-ng, Maltego e Shodan.

Foca (**Fingerprinting Organizations with Collected Archives**) è uno strumento OSINT utilizzato dai professionisti della sicurezza informatica per raccogliere dati da Internet. Può essere utilizzato per trovare informazioni su qualsiasi argomento, comprese persone, aziende e altre organizzazioni. Lo strumento raccoglie dati da una varietà di fonti come piattaforme di social media, siti Web e motori di ricerca. Lo strumento aiuta gli utenti a raccogliere informazioni in modo rapido ed efficiente fornendo loro una serie di strumenti per la ricerca, la raccolta e l'analisi dei dati raccolti. Fornisce agli utenti opzioni di filtro avanzate che consentono loro di restringere le ricerche e trovare facilmente le informazioni pertinenti. Foca ha anche funzionalità come l'analisi delle parole chiave che consente agli utenti di analizzare contenuti o immagini basati su testo al fine di identificare modelli o tendenze nei dati raccolti. Inoltre, offre altre funzionalità come la generazione automatica di report che consente agli utenti di generare rapidamente report senza dover raccogliere manualmente tutti i dati necessari.

Metagoofil è un potente strumento OSINT utilizzato per raccogliere informazioni pubblicamente disponibili su un particolare obiettivo. È particolarmente utile per i penetration tester, i professionisti della sicurezza e i ricercatori che hanno bisogno di raccogliere dati dai siti Web per eseguire ricognizioni sui loro obiettivi. Metagoofil è stato sviluppato da Edge Security nel 2006 come parte del framework per i suoi servizi di consulenza sulla sicurezza. Questo strumento può essere

utilizzato per eseguire la scansione di siti Web, motori di ricerca e archivi di documenti pubblici come PDF e documenti di Microsoft Office. Quindi cerca parole chiave specifiche relative al target e raccoglie le informazioni pertinenti da queste fonti. Con la sua interfaccia di facile utilizzo, Metagoofil consente agli utenti di trovare rapidamente file contenenti informazioni sensibili come nomi utente, password, indirizzi e-mail, indirizzi IP, ecc., che possono quindi essere utilizzati in ulteriori attacchi o progetti di ricerca.

GHunt è un nuovo strumento OSINT che consente agli utenti di estrarre informazioni da qualsiasi account Google utilizzando un'email. Le informazioni che GHunt estrae includono:

- ID Google,
- il nome del proprietario,
- foto pubbliche,
- modelli di telefoni,
- firmware dei telefoni,
- software installati,
- recensioni di Google Maps,
- possibile ubicazione fisica,
- possibile canale YouTube,
- possibili altri nomi utente,
- eventi da Google Calendar,
- se l'account è un bot di Hangouts,
- l'ultima volta che il profilo è stato modificato,
- servizi Google attivati (YouTube, Foto, Maps, News360, Hangouts, ecc.).

Yandex Images è la risposta russa al Google americano. Yandex è stato estremamente popolare in Russia e offre agli utenti la possibilità di cercare su Internet migliaia di immagini. Ciò è in aggiunta alla sua funzionalità di immagine inversa che è notevolmente simile a Google. Un'interessante funzionalità è la possibilità di ordinare le immagini per categoria, cosa che può rendere le ricerche più selettive e accurate.

N2YO.com permette di tracciare i satelliti da lontano, è un ottimo strumento per gli appassionati di spazio. Lo fa presentando un menù di satelliti ricercato regolarmente oltre a un database in cui è possibile effettuare query personalizzate lungo le linee di parametri come l'ID del comando spaziale, la data di lancio, il nome del satellite e un designatore internazionale. Permette anche di impostare avvisi personalizzati per conoscere gli eventi della stazione spaziale insieme a un live streaming della stazione spaziale internazionale (ISS).

TinEye è un motore per la ricerca inversa di immagini, cioè inviando un'immagine originale a TinEye è possibile ottenere tutte le informazioni richieste, come da dove proviene e come è stata utilizzata. Invece di utilizzare la corrispondenza delle parole chiave, utilizza una varietà di approcci per completare i suoi compiti, tra cui la corrispondenza delle immagini, la corrispondenza delle firme, l'identificazione della filigrana e numerosi altri database per abbinare l'immagine.

Have I Been Pwned è un servizio online che aiuta le persone a determinare se i loro dati personali sono stati compromessi. Funziona utilizzando gli indirizzi e-mail per tenere traccia delle violazioni dei dati, consentendo agli utenti di sapere se le loro informazioni sono state trapelate o rubate a

causa di un hack o di un altro incidente. Have I Been Pwned è stato creato nel 2013 da Troy Hunt, direttore regionale di Microsoft ed esperto di sicurezza. Il sito fornisce agli utenti informazioni dettagliate sulla fonte di eventuali violazioni che interessano i loro dati personali, nonché sui tipi di dati che potrebbero essere stati divulgati. Ciò consente loro di adottare le misure appropriate per proteggersi da attacchi futuri. Have I Been Pwned o HIBP tiene attualmente traccia di oltre 12 miliardi di account in oltre 600 gravi violazioni dei dati, fornendo uno dei database più completi per verificare se i dettagli del tuo account sono stati esposti online.

Tecniche di analisi di dati proveniente da fonti aperte basate su AI

Per rilevare potenziali minacce o attacchi verso un sistema, i moderni strumenti di sicurezza basati su AI monitorano un gran numero di eventi, in cerca di attività insolite o riconducibili a conseguenze dannose. Una volta individuato un rischio alla sicurezza, ad esempio lo sfruttamento di una vulnerabilità, il sistema può generare un **alert**, ovvero un messaggio relativo ad un evento di sicurezza che contiene informazioni sulle attività sospette rilevate. Un intrusion detection system (IDS) è un esempio di tool in grado di generare alert per eventi ritenuti sospetti. In questo contesto, è cruciale riuscire ad indentificare la correlazione tra alert diversi, così da avere una migliore comprensione dello sviluppo dell'attacco e poterne determinare la fonte e lo scopo. In alcuni tipi di attacco, come gli Advanced Persistent Threats, sono previste più fasi consecutive, correlate tra loro, che portano all'intrusione nel sistema bersaglio. Gli strumenti di analisi predittiva sfruttano tecniche di correlazione degli alert (sia dati storici che eventi in tempo reale) per rilevare automaticamente gli eventi anomali e quindi prevenire gli attacchi informatici nelle fasi iniziali.
Correlare i diversi eventi di sicurezza prevede innanzitutto la creazione di un contesto tra alert indipendenti. Questo può essere fatto in modi diversi, dando origine a tre principali categorie di metodi di correlazione: **similarity-based**, **step-based**, e **mixed**. I metodi similarity-based confrontano più eventi in base alla somiglianza dei loro attributi, che viene a sua volta calcolata utilizzando, ad esempio, funzioni di distanza euclidea, Mahalanobis o Manhattan. L'idea alla base di questo tipo di correlazione è che un gruppo di eventi simili può corrispondere allo stesso tipo di attacco. I metodi di correlazione step-based creano, invece, catene di eventi che ricostruiscono le azioni di un utente e analizzano le connessioni tra loro. Seguendo questo approccio è possibile sia abbinare gli eventi di sicurezza in base a specifici pattern (sfruttando basi di conoscenza delle vulnerabilità), sia definire catene di eventi in base alle loro relazioni statistiche e senza conoscenze pregresse. I metodi step-based, a loro volta, possono essere suddivisi in metodi **causal-based** e metodi di **data mining**. I primi analizzano la struttura causale degli eventi e producono una sequenza in cui ogni passo è determinato da quelli precedenti. I secondi, basati su data mining, cercano pattern nel dataset degli eventi tramite analisi statistiche. Infine, i metodi di tipo mixed utilizzano una combinazione di diversi algoritmi di correlazione, senza una evidente predominanza di uno sull'altro.
Un altro criterio che è possibile usare per caratterizzare le tecniche di correlazione di eventi è dettato dal tipo di rappresentazione scelta per rappresentare le informazioni (knowledge) a disposizione. In [46] vengono individuati quattro macro-modelli principali:

- modelli basati su regole (insieme di condizioni per confrontare e aggregare eventi di sicurezza);
- modelli semantici (linguaggi con sintassi e semantiche specifiche);
- modelli grafici (reti di nodi e archi);

- modelli di apprendimento automatico (raccolte di caratteristiche degli eventi e dei loro valori).

Nelle sezioni seguenti, forniamo esempi di implementazione di varie tecniche facenti parte delle diverse categorie elencate.

Modelli basati su regole

I modelli di correlazione basati su regole utilizzano spesso la conoscenza sulle relazioni causali degli eventi di sicurezza, presentate sotto forma di frasi condizionali. In questo caso, gli approcci alla correlazione degli eventi utilizzano spesso una base di conoscenza che contiene regole per la corrispondenza degli attributi degli eventi. Distinguiamo due sottocategorie di regole: regole di similarità e regole causali.
Le **regole di similarità** descrivono le condizioni per la somiglianza degli eventi in termini di attributi (features) e spesso utilizzano i valori di soglia dei coefficienti di correlazione e le misure di similarità. La soglia può essere impostata come un valore fisso del coefficiente di correlazione o calcolata dal valore medio della correlazione delle features [47]. L'approccio alla similarità degli eventi di sicurezza può anche tenere conto dei pesi degli attributi degli eventi assegnati in base alla classe di attacco [48]. Inoltre, la somiglianza può essere definita sia tra attributi di eventi dello stesso tipo, sia tra attributi di tipi diversi [49, 50].
Le **regole causali** descrivono le condizioni per la relazione causale degli eventi e utilizzano nozioni di *prerequisito* e *conseguenza*. In particolare, questi modelli collegano gli eventi in modo tale che le conseguenze degli eventi iniziali coincidano con i prerequisiti degli eventi successivi. La conoscenza dei prerequisiti e delle conseguenze può essere rappresentata anche sotto forma di *codebook* a matrice binaria, dove “1” indica la presenza di una relazione causale tra gli eventi e “0” la sua assenza. In RACC [51] (Real-time Alert Correlation based on Codebooks) i codebook corrispondono a scenari di attacco che vengono mappati agli avvisi in arrivo utilizzando operazioni matriciali. TempoCode-IoT [52] utilizza una rappresentazione della funzione di flusso basata sull'apprendimento non supervisionato di un codebook temporale che cattura i modelli chiave nei dati attraverso diverse finestre temporali. I centri dei cluster di ogni finestra temporale sono memorizzati come parole chiave.
Le due sottocategorie di regole sopra elencate possono essere combinate in regole composite. L'approccio di [53] utilizza il clustering di propagazione dell'affinità dei dati, identificando avvisi simili, e poi un metodo di prerequisiti e conseguenze per recuperare l'intero processo di attacco nelle reti IoT. Tradizionalmente, l'analisi degli eventi di sicurezza è fornita di default da regole esperte, come quelle fornite da sistemi SIEM come Open Source SIEM o Sigma, e regole programmate, ad esempio negli IDS Bro. I metodi di intelligenza artificiale consentono di estrarre automaticamente le regole di correlazione e di ridurre il costo delle specifiche manuali. In questo modo, la correlazione può essere costruita senza richiedere conoscenze predeterminate, in modo che il sistema permetta di trovare nuove correlazioni tra gli eventi. In ABE [54] (Automaton Based Engine) le regole di correlazione sono rappresentate da un albero di correlazione basato sull'analisi dei dati storici. In questo albero, i nodi sono operatori di tipo AND e OR e le foglie sono le azioni dell'attaccante. L'albero di correlazione viene poi trasformato in un automa in grado di riconoscere sequenze di eventi di sicurezza.
Esempi di event rule mining sono presentati in diverse pubblicazioni. SIRUS [55] (Stable and Interpretable RUle Set) estrae regole interpretabili da un classificatore random forest, cercando pattern frequenti negli alberi. L'algoritmo Case-crossover APriori [56] fornisce regole di associazione e causali che spiegano il verificarsi di eventi alluvionali. L'estrazione delle regole di associazione può

basarsi sulle caratteristiche temporali degli eventi, quando la segmentazione degli eventi viene eseguita utilizzando una finestra scorrevole. Questa analisi si basa sul calcolo della frequenza degli attributi in base alla soglia minima di supporto [57].

Modelli semantici

I modelli semantici utilizzano linguaggi con regole, sintassi e semantiche specifiche per stabilire una relazione tra input e output. In questo caso, gli eventi possono essere rappresentati da sequenze di caratteri che possono essere considerate come "parole" di un linguaggio formale, specificato da una grammatica formale. Nell'ambito della cybersecurity, un linguaggio di questo tipo consente di descrivere gli attacchi ad un sistema come sequenze di azioni che un agente malevolo compie per comprometterlo. Esempi di tali linguaggi sono STATL [58] (State Transition Analysis Technique Language) o ASTD [59] (Algebraic State Transition Diagram), che rappresentano una sequenza di eventi sotto forma di macchine a stati con azioni e variabili di stato. Altri linguaggi come SHEDEL [60] (Simple Hierarchical Event Description Language) e EDL [61] (Event Description Language) introducono una rete di Petri colorata in cui i nodi sono gli stati del sistema e le transizioni sono gli eventi correnti del sistema.
Le transizioni tra gli stati del sistema negli scenari di attacco possono essere definite anche utilizzando un linguaggio dichiarativo fuzzy. In una macchina a stati fuzzy, gli eventi sono rappresentati come insiemi fuzzy e le transizioni da uno stato all'altro sono anche descritte da una funzione di transizione fuzzy [62]. I possibili valori di ingresso sono determinati dai tipi di attacchi sotto forma di insiemi di eventi, mentre le transizioni da uno stato all'altro sono descritte da regole fuzzy.
I modelli semantici di correlazione degli eventi impiegano spesso tecniche di elaborazione del linguaggio naturale. Ad esempio, l'**event embedding** è una tecnica simile al word embedding, che interpreta i log testuali degli eventi come fossero insiemi di parole. Le parole (eventi) con lo stesso significato avranno una rappresentazione simile. I modelli di word embedding più utilizzati sono Word2Vec [63, 64] e GloVe [65] entrambi basati su tecniche di apprendimento non supervisionato. LogEvent2vec [66] e LogUAD [67] utilizzano Word2Vec per generare vettori di parole e costruire vettori pesati di features della sequenza di log. Doc2Vec [68] è simile all'algoritmo Word2vec, ma invece di vettorializzare le parole, crea un embedding vettoriale di frammenti di testo. Per l'approccio di [69] il corpus di addestramento è composto dagli eventi grezzi dei registri di sicurezza e ogni riga viene trattata come un paragrafo nell'addestramento del modello Doc2Vec. Log Transfer [70] rappresenta ogni modello di registro degli eventi utilizzando Glove, che tiene conto sia della corrispondenza globale delle parole che delle informazioni locali sul contesto. In questo modo, la presentazione dei modelli minimizza l'impatto dell'ordine delle parole (cioè della sintassi), preservando al contempo le informazioni semantiche. Questo aiuta a risolvere il problema che la sintassi dei log di sistemi di vario tipo è diversa, mentre la semantica dei log deve essere la stessa.
Il **sentence embedding** è simile al word embedding, ma invece delle parole, codifica l'intera frase in una rappresentazione vettoriale. Alcuni tra i più moderni modelli di sentence embedding sono ELMo [71] e *BERT* [72]. Questi modelli creano rappresentazioni di una parola sensibili al contesto, invece di creare un valore per ogni parola. *ELMo* (Embeddings from Language Model) considera il contesto in cui le parole vengono utilizzate, invece di creare un dizionario di parole. Uno degli esperimenti di [73] confronta le differenze tra l'embedding Word2Vec e la rappresentazione ELMo nella rete di predizione. La rappresentazione ELMo viene utilizzata per definire la matrice di embedding di ogni evento. I risultati mostrano che la rappresentazione ELMo può illustrare meglio le dinamiche contestuali e temporali nella predizione degli eventi. BERT (Bidirectional Encoder Representations

from Transformers) viene utilizzato attivamente non solo nell'elaborazione di linguaggi naturali, ma anche per quelli sintetici, come HTTP/HTTPS, per il rilevamento di attacchi nel traffico di rete [74].
Un altro linguaggio formale comune per descrivere gli eventi è il **linguaggio ontologico**. Possiamo rappresentare un'ontologia formale di eventi come Ω = (E,C,F,R,D), dove E è l'istanza di un evento, C è l'insieme dei concetti (ad esempio, tipi di eventi), F è l'insieme delle proprietà, R è l'insieme delle relazioni e D è il dominio dell'ontologia. L'apprendimento automatico di ontologie consente di creare automaticamente o semi-automaticamente ontologie estraendo termini per descrivere gli eventi. Durante il rilevamento degli eventi, il sistema cerca di estrarre relazioni complesse dalla sequenza di eventi, rappresentandoli come testo in linguaggio naturale. Esempi di strumenti linguistici di questo tipo sono ZSEE [75] (Zero-Shot transfer learning for Event Extraction) e OntoED [76].
Nell'ambito della sicurezza, l'ontologia può essere basata su una gerarchia di concetti che determinano le azioni degli aggressori per attuare attacchi di varie classi con vari gradi di dettaglio. L'ontologia delle intrusioni di [77] si basa sull'Intrusion Detection Messaging Format (IDMEF), un modello di dati per rappresentare le informazioni esportate da un IDS. Un'ontologia per il rilevamento degli attacchi può essere creata utilizzando reti neurali per apprendere l'embedding del testo come rappresentazione dei registri di eventi di sicurezza [78]. I modelli di attacco per l'ontologia di [79] sono estratti dai set di dati normalizzati utilizzando un algoritmo di Frequent-Item Mining basato sull'induzione orientata agli attributi (AOI-FIM). Questo algoritmo include l'aggregazione di eventi e la ricerca di pattern utilizzando il data mining.

Modelli grafici

I modelli grafici consentono di rappresentare la conoscenza degli eventi sotto forma di reti costituite da nodi (gli eventi) ed archi (che descrivono le relazioni tra gli eventi). I modelli grafici utilizzano spesso grafi orientati della forma G = (E,R,w), dove E è un insieme di nodi, R è un insieme di relazioni binarie su ExE, e la funzione w assegna un peso ad ogni arco. Questo modello viene anche chiamato "grafo di attacco" quando gli eventi rappresentano le varie fasi di un attacco alla sicurezza.
Esistono metodi automatizzati per costruire grafici basati sui dati degli eventi. NoDoze [80], OmegaLog [81], UNICORN [82] HOLMES [83] e WATSON [84] sono alcuni esempi di strumenti che analizzano le informazioni semantiche dei log e modellano i grafi di provenienza della conoscenza degli eventi.

- **NoDoze** si basa sull'idea che ogni evento sul grafo di provenienza è tanto più sospetto quanto sono sospetti gli eventi vicini sul grafo. In particolare, NoDoze crea un database di frequenza degli eventi e poi aggrega il punteggio di anomalia degli eventi vicini nel grafo.
- **OmegaLog** esegue un'analisi statica delle stringhe di messaggi di log e determina le loro relazioni temporali, creando un insieme di tutti i percorsi validi del flusso di log che possono verificarsi in fase di esecuzione. Una volta analizzato l'attacco, OmegaLog può utilizzare i percorsi del flusso di log per analizzare le relazioni di causa-effetto nella successione di eventi.
- **UNICORN** crea un grafico a blocchi che rappresenta l'intera storia delle chiamate di sistema e costruisce un modello evolutivo normale del comportamento del sistema per rilevare azioni anomale.
- **HOLMES** confronta tattiche, metodi e procedure che possono essere utilizzati per eseguire ogni fase dell'APT e crea un grafico di alto livello che riassume le azioni dell'attaccante in tempo reale.

- **WATSON** è un sistema di rilevamento delle intrusioni basato su host. Basandosi su informazioni contestuali, WATSON astrae i comportamenti in forma di vettori numerici e fornisce una rappresentazione vettoriale della semantica del comportamento.

Modelli di apprendimento automatico

Per quanto riguarda la combinazione di OSINT e IA, le principali pubblicazioni descrivono l'uso di algoritmi di apprendimento automatico e di elaborazione del linguaggio naturale per gestire e analizzare l'elevato volume di informazioni presenti su Internet.
In passato, il compito dei responsabili dell'analisi e dell'esecuzione di OSINT era quello di trovare informazioni nascoste o difficili da reperire. Oggi, di fronte a volumi crescenti di informazioni, l'altra sfida è quella di trovare informazioni pertinenti [85].
L'AI viene utilizzata generalmente su due fronti diversi: l'esplorazione dei contenuti disponibili su Internet e la loro elaborazione/analisi.
Poiché l'OSINT implica la produzione di conoscenza da una grande quantità di dati, le tecniche di AI utilizzate per il *data mining* possono essere impiegate in maniera molto naturale. Ruolo significativo nell'ambito è ricoperto dai sistemi e gli algoritmi di elaborazione del linguaggio naturale che sono stati utilizzati anche per la strutturazione dei dati, la traduzione automatica, l'estrazione di informazioni e l'analisi aggiuntiva dei risultati [86, 87].
Una delle ricerche più estese e approfondite per le applicazioni di sistemi intelligenti e machine learning in ambito OSINT è presentata in [88] dove, in particolare viene fatto notare che fino al 2014 le pubblicazioni che combinano OSINT e AI si focalizzano sull'uso di NLP per l'estrazione di informazioni su privacy, traffico di esseri umani e cybercrime da social media e big data in generale, mentre a partire dal 2015 iniziano a diffondersi lavori che usano sentiment analysis, text e media mining come tecniche sostitutive alle tecniche di intelligence tradizionali sia per la soluzione dei problemi che per l'identificazione delle minacce. Di particolare interesse i lavori più recenti estratti da [Evangelista2020] e mostrati in Figura 8 che coprono il range 2017-2019.
La ricerca in [89] pone l'attenzione sull'automatizzazione, anche con tecniche di IA, del processo OSINT. In particolare pone l'attenzione sul problema di *name-entity resolution* per identificare gli attori coinvolti e predire le loro future attività nei social media, l'attribuzione delle minacce e la costruzione di grafi per comprendere le interazioni fra i vari attori. I lavori presentati sono interessanti ma le tecniche presentate risultano ormai obsolete.
Focalizzando invece l'attenzione sui lavori più recenti è importante citare il survey proposto da Chaudhary and Bansal [90] completamente dedicato all'estrazione di informazioni relative ad attività terroristiche. In particolare, gli autori pongono l'attenzione su come la diffusione di Internet abbia portato sia alla proliferazione di informazioni legate al terrorismo e in particolare alla propaganda terrorista, ma anche all'avanzamento degli studi per l'estrazione di conoscenza sulla propaganda terrorista e allo scambio di messaggi fra terroristi. L'obiettivo generale è fornire una concisa rassegna degli strumenti e delle tecniche più attuali per l'estrazione di dati relativi al terrorismo da fonti di informazione pubblicamente disponibili dai media online, purtroppo limitatamente a quelle di tipo testuale. La letteratura è stata sistematicamente allineata alle tre fasi dell'estrazione OSINT: acquisizione dei dati, arricchimento dei dati e inferenza della conoscenza. Di particolare interesse la rassegna di vari dataset, sia pubblici che a pagamento, dove sono state collezionate nel tempo le informazioni su molteplici attività terroristiche. Dal punto di vista dell'OSINT la rassegna è interessante perché indica anche il tipo di fonte dati utilizzata per la costruzione del dataset (articoli su vari media, archivi elettronici, libri, riviste, documenti legali

pubblici etc.) e in partciolare mostra come le tecniche basate su deep learning siano le più diffuse quando i dati trattati sono di tipo testuale. Per quanto riguarda invece l'inferenza della conoscenza il lavoro illustra alcune tecniche per: *detection of hidden communities*, *radicalized content and people*, *unrest in public due to terrorist attacks*, *active offenders*, *active forums and websites spreading terroristic propagandas.*

L'inferenza di conoscenza dai dati estratti può essere fatta con due metodologie, ovvero l'analisi descrittiva (che descrive ciò che è accaduto sulla base dei dati acquisiti) e l'Analisi predittiva (che descrive ciò che può accadere in futuro sulla base delle osservazioni passate).

L'analisi descrittiva può essere implementata attraverso tecniche di network analysis per individuare i modelli di formazione delle reti sui media online (per rilevare le comunità nascoste e le reti operative dei terroristi) e tecniche di *sentiment analysis* per analizzare l'impatto emotivo (per esempio individuare l'impatto del terrorismo in termini di radicalizzazione, estremismo e disordini tra il pubblico).

Per quanto riguarda invece l'analisi predittiva il lavoro propone una rassegna di tecniche per prevedere il comportamento futuro delle organizzazioni terroristiche come ad esempio la predizione dell'autore di un attacco o la predizione delle armi di attacco.

Per quanto riguarda invece la citazione di articoli recenti proponenti tecniche per la soluzione di specifici problemi è interessante segnalare [91] e [92].

In [91] gli autori propongono un sistema di generazione di avvisi basato su *Twitter* per la segnalazione di nuovi argomenti rilevanti per la sicurezza informatica. Il sistema applica un classificatore supervisionato, basato sull'apprendimento attivo, che rileva i tweet contenenti informazioni rilevanti. L'approccio proposto, in particolare, riduce il numero di account e di tweet necessari per l'addestramento del classificatore, rendendo lo strumento adattabile al contesto specifico e favorendo la minimizzazione dei dati per l'Open Source Intelligence (OSINT). I tweet rilevanti sono infine raggruppati con un algoritmo di clustering greedy stream per identificare gli eventi significativi.

In [92] gli autori propongono un dataset e un word embedding model, creato a partire da BERT, specificamente adattati al dominio della cybersecurity. Tali strumenti possono chiaramente servire come elementi di base per i sistemi OSINT che sfruttano risorse testuali. Sia il dataset che il modello sono dichiarati pubblicamente disponibili ma non viene citato alcun riferimento a repository o pagine web (né sembrano essere reperibili da una ricerca su Google).

Title of publication	Applications area	Authors	Year
An investigation of using classification techniques in prediction of type of targets in Cyber attacks	Cybersecurity	Sina Pournouri, Shahrzad Zargari, Babak Akhgar	2019
BlackWidow: Monitoring the Dark Web for Cyber Security Information	Cybersecurity	Matthias Schafer, Markus Fuchs, Martin Strohmeier, Markus Engel, Marc Liechti, Vincent Lenders	2019
Cognitive security: A comprehensive study of cognitive science in cybersecurity	Cybersecurity	Roberto O Andrade, Sang Guun Yoo Facultad	2019
Design of a Classification Model for a Twitter-based Streaming Threat Monitor	Cybersecurity	Fernando Alves, Pedro M. Ferreira, Alysson Bessani	2019
Developing insights from social media using semantic lexical chains to mine short text structures	Social Media	Cecil Eng Huang Chua, Veda C. Storey, Xiaolin Li, Mala Kaul	2019
Enhancing Information Sharing and Visualization Capabilities in Security Data Analytic Platforms	Cybersecurity	Gustavo Gonzalez-Granadillo, Mario Faiella, Ibéria Medeiros, Rui Azevedo, Susana Gonzalez-Zarzosa	2019
Localising social network users and profiling their movement	Cybersecurity	Hector Pellet, Stavros Shiaeles, Stavros Stavrou	2019
Searching for Extremist Content Online Using the Dark Crawler and Sentiment Analysis	Military Purposes	Ryan Scrivens, Tiana Gaudette, Garth Davies, Richard Frank	2019
Turkish national cyber-firewall to mitigate countrywide cyber-attacks	Cybersecurity	Arif Sari	2019
A Supervised Machine Learning Based Approach for Automatically Extracting High-Level Threat Intelligence from Unstructured Sources	Cybersecurity	Yumna Ghazi, Zahid Anwar, Rafia Mumtaz, Shahzad Saleem and Ali Tahir.	2018
A survey on technical threat intelligence in the age of sophisticated cyber attacks	Cybersecurity	Wiem Tounsi, Helmi Rais	2018
Detecting Network Threats using OSINT Knowledge-based IDS	Cybersecurity	Ivo Vacas, Ibéria Medeiros, Nuno Neves	2018
Evaluating Automated Facial Age Estimation Techniques for Digital Forensics	Cybersecurity	Felix Anda, David Lillis, Nhien-An Le-Khac, Mark Scanlon	2018
Impact of AnonStalk (Anonymous Stalking) on users of Social Media: a Case Study	Cybersecurity	V. Kanakaris, K. Tzovelekis and D. V. Bandekas	2018
Is quantum computing becoming relevant to cyber-security?	Cybersecurity	Keegan Keplinger	2018
Managing cyber threat intelligence in a graph database	Cybersecurity	Seulgi Lee, Hyeisun Cho, Nakhyun Kim, Byungik Kim, Junhyung Park	2018
Modeling The Causes Of Terrorism From Media News: An Innovative Framework Connecting Impactful Events With Terror Incidents	Military Purposes	Truong Son Pham, Tuan-Hao Hoang	2018
Ontology population for open-source intelligence: A GATE-based solution	Languages and Translations	Giulio Ganino, Domenico Lembo, Massimo Mecella, Federico Scafoglieri	2018

Figura 8: Elenco lavori estratti da [Evangelista2020] limitatamente al periodo 2017-2019

Open source intelligence (OSINT) as support of cybersecurity operations. "Use of OSINT in a colombian context and sentiment Analysis"	Cybersecurity	Ricardo Andrés Pinto Rico, Martin José Hernández Medina, Cristian Camilo Pinzón Hernández, Daniel Orlando Díaz López, Juan Carlos Camilo García Ruíz	2018
Security OSIF: Toward Automatic Discovery and Analysis of Event Based Cyber Threat Intelligence	Cybersecurity	Ke Li, Hui Wen, Hong Li, Hongsong Zhu, Limin Sun	2018
Using Deep Neural Networks to Translate Multi-lingual Threat Intelligence	Languages and Translations	Priyanka Ranade, Sudip Mittal, Anupam Joshi and Karuna Joshi	2018
Applying fuzzy logic for sentiment analysis of social media network data in marketing	Social Media	Karen Howellsa, Ahmet Ertugan	2017
Classification of Colloquial Arabic Tweets in real- time to detect high-risk floods	Languages and Translations	Waleed Alabbas, Haider M. al-Khateeb, Ali Mansour, Gregory Epiphaniou, Ingo Frommholz	2017
Cloud security issues and challenges: a survey	Cybersecurity	Ashish Singh, Kakali Chatterjee	2017
Extracting Cyber Threat Intelligence From Hacker Forums: Support Vector Machines versus Convolutional Neural Networks	Cybersecurity	Isuf Deliu, Carl Leichter, Katrin Franke	2017
Toward a breakthrough Speaker Identification approach for Law Enforcement Agencies: SIIP	Languages and Translations	Khaled Khelif, Yann Mombrun, Gerhard Backfried, Farhan Sahito, Luca Scarpato, Petr Motlicek, Srikanth Madikeri, Damien Kelly, Gideon Hazzani, Emmanouil Chatzigavriil	2017
Utility and potential of rapid epidemic intelligence from internet-based sources	Social Media	S.J. Yan, A.A. Chughtai, C.R. Macintyre	2017

Figura 9: (Continua) Elenco lavori estratti da [Evangelista2020] limitatamente al periodo 2017-2019

I modelli di apprendimento automatico, come lo *shallow* e il *deep learning*, utilizzano i frame come rappresentazione dei dati. Un frame è una struttura di dati che include una collezione di attributi e valori e consiste in una collezione di slot e valori di slot di qualsiasi tipo e dimensione. Questa struttura consente di utilizzare grandi quantità di informazioni sugli eventi e di analizzarle con metodi di AI come l'analisi dei cluster e l'apprendimento automatico.
Quando esiste una base di conoscenza contenente tutti gli eventi normali e anormali e i loro attributi, il problema del rilevamento degli eventi e della previsione delle sequenze di eventi può essere ridotto ad un problema di classificazione multiclasse. Il rilevamento di anomalie nella sequenza di eventi è considerato un compito di apprendimento non supervisionato o semi-supervisionato. In entrambi i casi, si distinguono le seguenti fasi principali di risoluzione del problema:

1. identificazione delle features informative degli eventi;
2. selezione e addestramento di un modello (algoritmo) in grado di assegnare la sequenza di eventi corrente a una determinata classe;
3. calcolo dell'affidabilità e dell'accuratezza.

Di norma, tutte e tre le fasi vengono ripetute iterativamente fino a quando non vengono individuati un insieme di caratteristiche e un modello che soddisfino i criteri di affidabilità e accuratezza specificati.
I modelli di **shallow learning** sono tecniche tradizionali di apprendimento automatico che possono essere utilizzate per la correlazione degli avvisi attraverso la mappatura di features quali i valori degli attributi, le frequenze degli eventi, ecc. Il numero di avvisi giornalieri, la frequenza di occorrenza degli eventi e le funzioni relazionali ottenute dall'analisi del grafo sociale sono utilizzate come features per l'addestramento. In [93] profilano i dati del malware per estrarre gli scenari di

attacco utilizzando gli algoritmi k nearest neighbor (k-NN), decision tree (DT) e support vector machine (SVM). Per il monitoraggio dei Big Data nei sistemi IoT, l'approccio di [94] utilizza una struttura che utilizza principal component analysis (PCA), DT, SVM, k-NN, gaussian näive Bayes (GNB) e artificial neural network (ANN). Il modello SMOTE-RF [95] combina gli algoritmi SMOTE e random forest (RF) per risolvere il problema della classificazione sbilanciata e della multiclassificazione nei dataset APT. SMOTE aumenta il numero di campioni di minoranza attraverso l'interpolazione di k-NN per migliorare la distribuzione nei set di dati sbilanciati. Successivamente, viene eseguito l'apprendimento multiclasse basato su RF.

Una **Recurrent Neural Network** (RNN) consente l'analisi di dati sequenziali, come serie temporali o testi in linguaggio naturale. Approcci come Tiresias [96] DeepLog [97], OC4Seq [98] utilizzano RNN per prevedere eventi futuri sulla base di osservazioni precedenti per tracciare comportamenti anomali. A partire da eventi storici osservati, Tiresias calcola una distribuzione di probabilità di possibili eventi per prevedere i passi specifici che saranno compiuti da un possibile attaccante. Analogamente, OC4Seq utilizza Gated Recurrent Units (GRU) per il rilevamento di anomalie nelle sequenze di eventi. Infine, DeepLog utilizza la Long Short-Term Memory (LSTM) e apprende le correlazioni e i modelli incorporati in una sequenza di voci di registro prodotte dai normali percorsi di esecuzione del sistema.

Le **Convolutional Neural Networks** (CNNs) vengono spesso utilizzate per elaborare matrici di dati alla ricerca di modelli specifici. Ad esempio, [99] codificano la sequenza di chiamate di sistema in una "immagine" bidimensionale di lunghezza fissa, molto adatta all'analisi della CNN. A tale scopo, viene utilizzato N-Gram, la cui idea principale consiste nel tagliare la finestra di osservazione e ottenere i segmenti di sequenza di lunghezza esattamente N. Le CNNs vengono anche utilizzate in DeepCorr (Nasr et al, 2018) per studiare la funzione di correlazione del flusso nella rete Tor.

Un **autoencoder** è una rete neurale simmetrica usata per studiare le features degli eventi tramite approcci non supervisionati. [100] utilizza un modello di deep autoencoder per estrarre automaticamente le caratteristiche degli eventi e codificare i vettori degli attacchi APT. Eventuali problemi di sovra-generalizzazione possono essere risolti attraverso Network Anomaly Detection [101] con MemAE (Memory-augmented deep Auto-Encoder) che consentono di ottenere buoni risultati anche quando il campionamento normale e quello degli attacchi hanno features comuni.

La **Generative Adversarial Network** (GAN) utilizza un meccanismo basato su avversario per estrarre relazioni implicite tra gli eventi. [102] usano un approccio a due reti: una rete insegnante, che codifica i dati degli eventi in rappresentazioni vettoriali della conoscenza per l'apprendimento delle features e la rete studente che elabora testi grezzi per il rilevamento degli eventi. Questo approccio non richiede toolkit aggiuntivi, eliminando alla radice il problema di propagazione degli errori riscontrato negli approcci a pipeline.

I modelli **ensemble** combinano diversi modelli di apprendimento, sia supervisionati che non supervisionati, shallow o deep. L'approccio di [103] utilizza un insieme di modelli random forest, regressione logistica e autoencoder per rilevare e prevedere le interruzioni della rete mobile utilizzando più set di dati riguardanti l'attività dell'utente. [104] descrivono un predittore di tipo ensemble che utilizza una deep neural network (DNN) e un modello di regressione lineare per rilevare letture anomale di sensori cyber-fisici. In questo sistema, la misurazione di ciascun sensore può essere prevista in funzione degli altri sensori. Gli studi sulla combinazione di modelli spesso comportano la valutazione e la determinazione delle combinazioni ottimali di modelli di intelligenza artificiale e dei loro parametri per ottenere risultati nel modo più efficace. [105] utilizzano tre modelli di classificazione per rilevare e predire gli attacchi APT. Inoltre, una combinazione di modelli CNN e LSTM viene spesso utilizzata per rilevare e prevedere APT [106, 107]. Il vantaggio del modello CNN-LSTM nell'analisi degli eventi di sicurezza è che tale architettura funziona bene con attività in cui i dati grezzi hanno una struttura esplicita e proprietà temporali.

Modelli ibridi

Alcuni approcci e sistemi combinano diversi metodi di correlazione degli eventi. Di norma, i sistemi che utilizzano una combinazione di modelli basati su similarità e regole casuali presumono che gli eventi più simili possano essere associati allo stesso scenario di attacco. La ricostruzione dello scenario di attacco consiste in tre fasi principali:

1. identificare gli eventi di sicurezza correlati;
2. abbinare un sottoinsieme di eventi allo scenario appropriato;
3. ordinare la sequenza di eventi.

La correlazione può basarsi sulla similarità degli eventi in base a diversi parametri (ad esempio, indirizzi IP e porte di origine e destinazione) e lo scenario può essere rappresentato come un grafo [108,109]. Un'altra modalità di correlazione è l'analisi di eventi semanticamente simili. MAAC [110] (Multi-step Attack detection by Alert Correlation) utilizza Doc2vec per ottenere la rappresentazione semantica di un alert e calcola la distanza del coseno del vettore generato. L'approccio di [111] utilizza il modello Word2Vec per convertire gli alert in valori continui a bassa dimensione e abbinare quelli semanticamente simili. Il metodo Log2Vec [112] utilizza Log-Specific Word Embedding (LSWE) Word2Vec per la rappresentazione delle parole che migliora le informazioni semantiche e relazionali specifiche del dominio. LSWE utilizza due metodi di embedding delle parole: Lexical information Word Embedding (LWE) e Semantic Word Embedding (SWE). LWE predice la parola target in modo che la sua distanza nella rappresentazione vettoriale sia la più vicina possibile ai suoi sinonimi e la più lontana possibile dai suoi contrari. SWE, invece, definisce relazioni associative tra le parole.
L'estrazione degli attributi degli eventi di sicurezza può essere eseguita anche con metodi di apprendimento automatico. Il sistema MIF [113] (Multi-Information Fusion system) estrae i flussi di alert anomali utilizzando una CNN chiamata Convolution and agent decision Tree network (CTnet) e ricostruisce poi lo scenario di attacco utilizzando un modulo di fusione basato su grafi. CTnet valuta i rischi di attacco che, insieme alle informazioni sui nodi di attacco e sul tempo di attacco, vengono utilizzati per costruire un modulo di fusione. La catena di attacchi ad alto rischio viene recuperata utilizzando un algoritmo Time-Weighted Depth-First Search (TW-DFS). Le informazioni sul peso determinano il percorso attraverso i nodi a maggior rischio di attacco, mentre le informazioni sul tempo aiutano a rimuovere le correlazioni non temporali degli attacchi.
Recentemente, i metodi basati sull'intelligenza artificiale sono spesso utilizzati per analizzare grafi di eventi. REGNN [114] (Real-time Event Graph Neural Network) può prevedere eventi in tempo reale costruendo grafi dinamici eterogenei. Questo modello crea grafi di provenienza degli eventi basati sul comportamento degli utenti e poi utilizza reti neurali ricorrenti per modellare la dipendenza temporale degli eventi passati e incorporare eventi in tempo reale. A sua volta, la CNN può essere applicata ad una rete convoluzionale a grafo (GCN) [115], in cui il vettore di convoluzione per ogni nodo è calcolato dai vettori di rappresentazione più vicini. Per rilevare e prevedere gli eventi, la GNN utilizza il vettore evento corrente nel grafo. Il GDN [116] (approccio basato su Graphical Deviation Network) esamina il grafo delle relazioni tra i sensori e rileva le deviazioni da questi schemi. Questo approccio consente di rilevare le anomalie senza dati preliminari sulla struttura dei grafi.

Tecniche e indicatori di misurazione del rischio di corruzione da fonti aperte

Premessa

Misurare la corruzione è tutt'oggi una sfida in larga parte aperta per diverse ragioni, legate in primo luogo al fatto che la corruzione è, per sua natura, un'attività sommersa. Come per altre attività illecite, gli attori coinvolti hanno interesse a occultare o falsificare le informazioni sulle proprie attività, a nasconderne o distruggerne le prove. Le attività corruttive quindi, specialmente quelle che si realizzano nella forma di business economico, non si osservano mai direttamente nei dati.
Per di più, la corruzione è un fenomeno intrinsecamente complesso poiché comprende diverse attività a livelli diversi di gravità, da quelle triviali a quelle gravi. Esistono inoltre diversi tipi di corruzione, a seconda dei settori coinvolti (privati o pubblici; politici o amministrativi), degli attori coinvolti (ufficiali pubblici o privati, cittadini, politici) e del grado di formalizzazione degli atti corruttivi (sistematici o occasionali). Questa complessità si traduce in una generale difficoltà di sviluppare una definizione di corruzione singola e onnicomprensiva, e questo a sua volta conduce a problemi generali di validità degli strumenti di misurazione della corruzione.
Esistono diversi strumenti e indicatori per misurare la corruzione e il rischio che si verifichi. Ciascuno presenta limiti e vantaggi specifici, tali da renderli preferibili ad altri solo in determinati contesti e per rispondere ad obiettivi specifici e ben determinati. In termini assoluti, e sotto il profilo strettamente misuratorio, nessuno può essere considerato superiore e dunque preferibile agli altri.
Tuttavia, nell'ultimo decennio, il ricorso a indicatori di rischio corruttivo nel settore degli appalti pubblici (i.e., red flags di rischio di corruzione) ha assunto sempre maggior rilievo in ragione, prima di tutto, del peculiare peso del fenomeno corruttivo su tale mercato. L'ambito degli appalti pubblici, il processo attraverso il quale le autorità pubbliche acquistano lavori, beni o servizi, rappresenta uno dei più vulnerabili alla corruzione. I rischi di corruzione possono verificarsi in ogni fase del processo e sono esacerbati dal volume delle transazioni economiche e degli interessi finanziari in gioco. In effetti, gli appalti pubblici rappresentano oltre il 14% del PIL dell'UE. In Italia, gli appalti pubblici rappresentano oltre il 10% del PIL e oltre il 21% della spesa pubblica generale. La corruzione nel settore degli appalti ha quindi enormi costi diretti – tra cui perdita di fondi pubblici attraverso aggiudicazioni improprie, maggiori spese e minore qualità di beni, servizi e lavori e costi indiretti – da distorsione della concorrenza, ad accesso limitato al mercato e ridotta propensione al business da parte degli investitori stranieri.
Anche l'Autorità Nazionale Anticorruzione (ANAC) ha scelto di concentrarsi su questo insieme di indicatori, anche in ragione della disponibilità della Banca Dati Nazionale dei Contratti Pubblici (BDNCP), una banca dati open, con molto ampia copertura longitudinale e che consente il calcolo degli indicatori con un estremo grado di dettaglio.

La corruzione come variabile latente e multidimensionale

L'incipit del contributo di G. Arbia contenuto nel volume Misurare la corruzione oggi: obiettivi, metodi, esperienze (2018) "Non sempre i fenomeni sono misurabili direttamente, ma sempre, come nel caso del vento, ne possiamo osservare gli effetti diretti e indiretti", nel ricordarci che la corruzione, come il vento, è un fenomeno non direttamente osservabile, individua la prima e più importante sfida che si incontra nella sua misurazione. Come per altre attività illecite, la corruzione è infatti invisibile all'esterno. Per sua natura, la corruzione è un fenomeno elusivo e sfuggente che idealmente non lascia alcuna traccia. Gli attori coinvolti hanno interesse a occultare o a falsificare

le informazioni sulle proprie attività. Normalmente, la corruzione è un business economico. Ma non ci sono quasi mai record, trasferimenti o bonifici bancari, ricevute o fatture che possano costituire l'oggetto tangibile e direttamente quantificabile del processo di misurazione.
Tuttavia, la circostanza che un fenomeno non sia direttamente osservabile non equivale alla impossibilità di misurarlo. Al contrario, nell'ambito delle scienze applicate, dalla statistica economica a quella sociale, ci si confronta spesso con il problema di misurare fenomeni non osservabili direttamente, e per questo definiti tecnicamente variabili latenti. Ne sono esempi i sentimenti, la fiducia dei consumatori, le abilità degli studenti, la qualità della vita, lo sviluppo di uno Stato e così via. Il problema della misurazione della corruzione non è quindi un caso isolato negli studi statistici. In tutti questi casi, poiché non è possibile osservare e misurare direttamente il fenomeno oggetto di interesse, si ricorre al calcolo di misure indirette (variabili proxy o indicatori) che approssimano il fenomeno non osservabile. Gli indicatori red flags di rischio di corruzione ne sono un esempio.
Per di più, la corruzione è anche un fenomeno intrinsecamente complesso poiché comprende diverse attività a livelli diversi di gravità, da quelle triviali a quelle gravi, dalla bustarella per evitare un'azione penale per una trasgressione del codice della strada, fino alla falsificazione di decisioni pubbliche per perseguire interessi privati illeciti. Esistono anche diversi tipi di corruzione, a seconda dei settori coinvolti (privati o pubblici; politici o amministrativi), degli attori coinvolti (ufficiali pubblici o privati, cittadini, politici), del grado di formalizzazione degli atti corruttivi (sistematici o occasionali).
Il problema della misurazione di fenomeni complessi e multidimensionali come la corruzione non è un caso isolato negli studi statistici in ambito applicato. Diversi fenomeni in ambito socio-economico sono, infatti, intrinsecamente complessi e la complessità implica multidimensionalità. Ad esempio, il livello di sviluppo di uno Stato è un obiettivo chiave complesso e multidimensionale. È stato considerato unidimensionale fino agli anni '60, quando l'approccio multidimensionale ha spostato il tradizionale focus sulla dimensione monetaria come proxy sufficiente del benessere umano. Così come non è possibile ridurre il livello di sviluppo di uno Stato ad una unica singola dimensione monetaria, allo stesso modo non si possono forzare altri costrutti multidimensionali come la qualità della vita, la qualità percepita di un bene o servizio, la stessa corruzione, in un unico valore numerico o indicatore composito perché rischieremmo di appiattire differenze significative tra unità e di operare una eccessiva semplificazione delle differenze tra sottodimensioni del fenomeno che non rispecchia la sua complessità.
Qualunque sia il terreno di applicazione, la necessità di misurazioni multidimensionali di fenomeni complessi emerge dunque da una serie di imperativi (politici, istituzionali, ecc.) in gran parte sovrapponibili tra campi applicativi e riconducibili alla necessità di identificare e trattenere le diversità a fini di policy making. Quando gli scienziati conducono studi sulla qualità della vita, la domanda di misurazioni multidimensionali deriva principalmente da un mandato per ridurre le disuguaglianze tra popolazioni in posizioni socio-economiche svantaggiate diversificate per la destinazione di fondi e interventi pubblici. In maniera piuttosto simile, quando l'obiettivo è studiare il rischio di corruzione è condizione cruciale mantenere la lente multidimensionale e disporre di strumenti adeguati e realistici necessari ad analizzare la diversa presenza e consistenza del fenomeno corruttivo, a evidenziare i diversi fattori di rischio e a monitorare l'impatto prodotto dalle misure di prevenzione e contrasto al fine ultimo di sostenere l'integrità di uno Stato sotto il profilo politico-sociale, giuridico ed economico.

Indicatori red flags di rischio di corruzione: la selezione operata dall'Autorità Nazionale Anticorruzione (ANAC) per misurare il rischio di corruzione a livello territoriale

Gli indicatori red flag sono al centro dei sistemi di valutazione del rischio di corruzione sviluppati per identificare i rischi di corruzione e definire strategie efficaci per mitigarli in una logica preventiva. Negli ultimi anni, si è assistito ad una maturazione e un consolidamento della ricerca nell'ambito della valutazione del rischio di corruzione negli appalti pubblici grazie soprattutto alla disponibilità di banche dati in formato leggibile e open e allo sviluppo di nuove tecnologie di data science basate sulla raccolta e l'elaborazione incrociata dei dati sugli appalti pubblici con altre fonti di dati detenuti dalle pubbliche amministrazioni.

Attualmente, la letteratura internazionale tende a convergere verso la selezione di una serie di indicatori di rischio di corruzione considerati più rilevanti. Tra questi, i più diffusi sono [124]: offerta singola, cioè una sola offerta ricevuta nell'ambito di una procedura di gara; mancata pubblicazione del bando di gara in gazzetta ufficiale (quando la pubblicazione non sia obbligatoria); lunghezza relativa dei criteri di ammissibilità; costo relativo della documentazione di gara; esclusione di tutti gli offerenti tranne uno; peso di criteri di valutazione non relativi al prezzo, ovvero proporzione tra criteri di valutazione non correlati al prezzo nel complesso di tutti i criteri di valutazione; procedura annullata rilanciata; durata del periodo di decisione, ovvero estensione temporale tra il termine di presentazione dell'offerta e l'aggiudicazione dell'appalto; modifica del contratto durante la fase di consegna; allungamento e proroghe dei contratti; incremento in itinere del valore economico del contratto.

Nell'ambito del progetto «Misurazione del rischio di corruzione a livello territoriale e promozione della trasparenza», finanziato dal Programma Operativo Nazionale Governance e Capacità istituzionale (2014-2020), l'Autorità Nazionale Anticorruzione (ANAC) ha selezionato e calcolato un insieme di indicatori di rischio corruzione a livello provinciale utili per sostenere la prevenzione e il contrasto all'illegalità e promuovere la trasparenza nell'azione della Pubblica Amministrazione. Gli indicatori selezionati sono indicati di seguito, insieme a una breve descrizione della motivazione della scelta.

1. Indicatore numero OEPV

L'offerta economicamente più vantaggiosa (OEPV) è un criterio di aggiudicazione del contratto di gara mediante il quale la stazione appaltante confronta le offerte con riguardo al miglior rapporto qualità/prezzo. Applicando il criterio dell'offerta economicamente più vantaggiosa, la stazione appaltante attribuisce punteggi per la qualità di lavoro/servizio/fornitura e per la voce prezzo, e aggiudica il contratto all'operatore economico che riceve il punteggio più alto nelle due voci.

Rileva la frazione di procedure aggiudicate utilizzando il criterio dell'offerta economicamente più vantaggiosa rispetto al totale. È uguale al rapporto tra:

- numeratore: numero procedure aggiudicate con il criterio dell'offerta economicamente più vantaggiosa;
- denominatore: numero totale procedure aggiudicate.

L'offerta economicamente più vantaggiosa, sebbene trovi anche con l'introduzione delle ultime direttive uno spazio sempre maggiore come criterio di scelta da utilizzare, presenta un più alto rischio di discrezionalità rispetto al criterio del prezzo più basso e per questo è considerato una proxy di rischio di corruzione [117, 118, 121]. Infatti, col criterio del prezzo più basso, la stazione appaltante confronta le offerte solo con riguardo al maggior ribasso di prezzo rispetto alla base d'asta. Aggiudica di conseguenza la gara all'operatore economico che offre il prezzo immediatamente più basso rispetto alla soglia. Diversamente, quando la stazione appaltante aggiudica un contratto col criterio dell'offerta economicamente più vantaggiosa, confronta le offerte con riguardo al miglior rapporto qualità/prezzo. Nella valutazione della qualità, possono entrare in gioco criteri discrezionali e difficilmente quantificabili come, ad esempio, la valutazione del pregio tecnico, delle caratteristiche estetiche e funzionali, degli aspetti di innovatività, delle condizioni di consegna e di esecuzione del servizio o lavoro, etc.

2. Indicatore sul numero delle procedure non aperte

Rileva la frazione di procedure non aperte (affidamenti diretti e procedure negoziate) rispetto al totale. È uguale al rapporto tra:

- numeratore: numero di procedure non aperte;
- denominatore: numero totale di procedure.

L'indicatore ha lo scopo di valutare la percentuale di procedure non aperte (procedure negoziate con o senza previa pubblicazione di un bando, affidamenti diretti, cottimi fiduciari, ecc.) sul totale delle procedure utilizzate da una medesima stazione appaltante in un determinato arco temporale.

L'indicatore di per sé non segnala illegittimità poiché è possibile che le procedure prescelte da una stazione appaltante diverse da quelle aperte o ristrette rispettino tutti i requisiti imposti dalla normativa vigente. Tuttavia, una elevata percentuale di aggiudicazioni affidate secondo meccanismi non concorrenziali insieme ad altri indicatori potrebbe segnalare una patologia da monitorare in maniera specifica [117,118,119,120, 121, 122, 123].

3. Indicatore sul valore delle procedure non aperte

È analogo al precedente, rilevando però la frazione del valore economico delle procedure non aperte (affidamenti diretti e procedure negoziate) sul valore totale delle procedure [117]. È uguale al rapporto tra:

- numeratore: somma del valore economico delle procedure non aperte;
- denominatore: somma del valore economico del totale delle procedure.

4. Indicatore del numero di contratti aggiudicati e modificati per effetto di almeno una variante

Rileva la frazione dei contratti che in fase di esecuzione sono stati interessati da almeno una variante in corso d'opera rispetto al totale delle procedure. È uguale al rapporto tra:

- numeratore: numero di procedure interessate da almeno una variante;
- denominatore: numero totale di procedure aggiudicate e concluse.

L'indicatore rappresenta una misura della proporzione di contratti che in fase di esecuzione sono interessati da varianti in corso d'opera. L'indicatore può effettivamente segnalare una patologia in determinate circostanze ad esempio nei casi di varianti necessarie a superare possibili errori progettuali [117,122,125].

5. Indicatore di scostamento dei costi di esecuzione

Rileva lo scostamento dei costi ed è la media aritmetica del rapporto (calcolato per ogni procedura) tra:

- numeratore: costo effettivo (importo finale a consuntivo);
- denominatore: costo preventivato (importo di aggiudicazione).

L'indicatore può essere utile a valutare eventuali comportamenti di "moral hazard" in corso di esecuzione del contratto. Può accadere, infatti, che alcuni operatori economici facciano ribassi molto forti in sede di aggiudicazione vincendo un determinato contratto per poi recuperare lo sconto dichiarato durante l'esecuzione. L'aumento dei costi di esecuzione rispetto a quelli inizialmente previsti può essere legato a circostanze impreviste ed imprevedibili ed essere pertanto giustificato ma potrebbe anche essere legato ad una connivenza tra operatore economico e stazione appaltante per aumentare artificiosamente i costi dell'appalto [117,122,123,125].

6. Indicatore di scostamento dei tempi di esecuzione

Rileva lo scostamento dei tempi di esecuzione ed è la media aritmetica del rapporto (calcolato per ogni procedura) tra:

- numeratore: durata effettiva;
- denominatore: durata prevista.

Analogamente all'indicatore sullo scostamento dei costi di esecuzione, l'indicatore di scostamento dei tempi ha la finalità di valutare comportamenti opportunistici da parte dell'operatore economico assecondati dalla stazione appaltante.

Anche questo indicatore, come il precedente, deve essere letto con prudenza. Infatti, eventuali scostamenti tra tempi di realizzazione previsti e tempi effettivi possono essere giustificati da sospensioni legittime [118,121,122,123,125].

7. Indicatore di inadempimento delle comunicazioni di aggiudicazione

Rileva la frazione di procedure per cui è avvenuta la comunicazione all'autorità della scheda di aggiudicazione rispetto al totale. È uguale al rapporto tra:

- numeratore: numero di procedure per cui è stata comunicata l'aggiudicazione;
- denominatore: numero di procedure la cui scheda di aggiudicazione deve essere comunicata.

L'inadempimento all'obbligo di comunicazione dei dati alla Banca Dati Nazionali dei Contratti Pubblici (nel caso di specie dell'aggiudicazione della procedura) è manifestazione di cattiva condotta delle amministrazioni, che potrebbe essere collegata a fattispecie di corruzione.

8. Indicatore di inadempimento delle comunicazioni di fine lavori

Rileva la frazione di procedure per cui è avvenuta la comunicazione all'autorità della scheda di fine lavori, rispetto al totale. È uguale al rapporto tra:

- numeratore: numero di procedure per cui è stata comunicata la scheda di fine lavori;
- denominatore: numero totale di procedure la cui scheda di fine lavori deve essere comunicata.

L'inadempimento all'obbligo di comunicazione dei dati alla Banca Dati Nazionali dei Contratti Pubblici (nel caso di specie della fine dei lavori) è manifestazione di cattiva condotta delle amministrazioni, che potrebbe essere collegata a fattispecie di corruzione.

9. Indicatore Offerta singola

Rileva la proporzione di procedure per le quali è stata presentata una sola offerta, da parte di un solo partecipante al bando, rispetto al totale delle procedure aggiudicate dalla stazione appaltante.

	Offerte ammesse		
Offerte presentate	**Una**	**Più di una**	
Una	n_{11}	-	
Più di una	n_{21}	n_{22}	
Totale	$n_{.1}$	$n_{.2}$	N

Figura 10: Descrizione del tipo di procedure in base alle offerte presentate e ammesse (N è il numero totale di procedure considerate)

Considerando la Figura 10, possiamo avere la seguente classificazione:

- procedure per cui è stata presentata una sola offerta, che poi è stata ammessa (n_{11});
- procedure per cui sono state presentate più offerte, ma una sola è stata ammessa (n_{21});
- procedure per cui sono state presentate più offerte, di cui ne sono state ammesse più di una (n_{22}).

La letteratura scientifica a cui si è fatto riferimento è quella, ampiamente consolidata a livello internazionale [119,121,123,124,125,127] che considera l'assenza di competizione come una condizione ad alto rischio di corruzione nella procedura di appalto. Infatti, l'elusione dei principi della concorrenza leale è associata ad una restrizione ingiustificata dell'accesso ai contratti. Il caso più evidente di difetto di competizione si verifica proprio quando una procedura di appalto riceve una sola offerta. Si tratta di una condizione che può consentire l'aggiudicazione del contratto a prezzi più alti di quelli di mercato e che può essere il riflesso di rapporti particolari tra stazione appaltante e aziende e/o di accordi preliminari tra esse.
È calcolato mediante il rapporto tra le seguenti quantità (si veda Figura 10):

- numeratore: numero di procedure con un'unica offerta presentata e ammessa (n_{11});
- denominatore: numero totale di procedure aggiudicate (N).

10. Indicatore proporzione di offerte escluse

Rileva la media aritmetica del rapporto tra:

- numeratore: numero offerte escluse (calcolate per ogni procedura);
- denominatore: numero offerte presentate.

La letteratura - si veda, a titolo di esempio: European Anti-Fraud Office [117,122,126,128,129,130] insiste in modo deciso sulla condizione di rischio che si verifica quando il numero di offerte in gara è limitato. Questa condizione può verificarsi non solo in caso di limitata presentazione di offerte da parte delle aziende ma anche a seguito di esclusione delle stesse in fase di valutazione da parte della stazione appaltante. L'indicatore intende misurare, in fase di valutazione delle offerte pervenute da parte della stazione appaltante, il peso di quelle escluse sul totale delle offerte ricevute. L'ipotesi di fondo è che il rischio di corruzione è tanto più alto quanto maggiore è la quota di offerte escluse. Infatti, l'esclusione di gran parte delle offerte potrebbe essere il segnale di una strategia di selezione da parte della stazione appaltante diretta a favorire le sole aziende ad essa collegate da rapporti di tipo particolaristico, respingendo quelle "indesiderate".

11. Indicatore Esclusione di tutte le offerte tranne una

Rileva la frazione di procedure per cui sono state escluse tutte le offerte tranne una (si veda Figura 10), rispetto al totale di procedure con una sola offerta ammessa. È uguale al rapporto tra

- numeratore: numero di procedure con una sola offerta ammessa a fronte di più offerte presentate (n_{21});
- denominatore: numero di procedure con una sola offerta ammessa ($n_{.1}$).

L'esclusione di tutte le offerte tranne una, esattamente come l'offerta singola, è un caso evidente di assenza di competizione leale. Quando tutte le offerte vengono escluse tranne una, così come quando viene presentata una sola offerta, non c'è competizione e il rischio di corruzione è considerato alto. La letteratura di riferimento per questo indicatore è la medesima indicata per l'indicatore Offerta singola.

12. Indicatore proporzione di offerte escluse in procedure con tutte escluse tranne una

Il calcolo di questo indicatore è simile a quello dell'indicatore 10, ma la media è calcolata considerando soltanto le procedure con una sola offerta ammessa a fronte di più offerte presentate (n_{21} in Figura 10).

L'indicatore è una variante dell'indicatore proporzione di offerte escluse e misura la proporzione di offerte escluse tra le sole procedure nelle quali sono state escluse tutte le offerte tranne una. La proposta di questo indicatore nasce dall'esigenza di distinguere i casi in cui il numero di offerte è elevato e l'esclusione finalizzata a lasciarne in gioco solo una riguarda un elevato numero di offerte, dai casi in cui il numero di offerte è ridotto e, quindi, l'esclusione riguarda un numero limitato di offerte. L'ipotesi è che il rischio di corruzione sia più alto nella prima delle due circostanze sopra delineate, ovvero quando, in fase di valutazione delle offerte pervenute da parte della stazione appaltante, quest'ultima ne esclude un numero importante, per lasciarne in gioco solo una.

13. Indicatore proporzione di contratti aggiudicati alla stessa azienda

L'indicatore parte da quello noto in letteratura come *Winner's share of issuer contracts*. Nella sua versione originaria [119,123,126] questo indicatore misura la proporzione di valore contrattuale aggiudicato (€) da una stazione appaltante ad un'azienda aggiudicataria rispetto al valore totale di tutti i contratti aggiudicati dalla medesima stazione appaltante. Una versione alternativa dello stesso, di seguito proposta, considera il numero di procedure anziché il valore economico aggiudicato. Questo indicatore, pertanto, valuta la ricorrenza o frequenza con cui una stazione appaltante aggiudica i propri contratti ad una stessa azienda. La letteratura - si vedano, a titolo di esempio, i policy report di organismi internazionali come World Bank Group (disponibili al link: https://www.worldbank.org/en/about/unit/integrity-vice-presidency) - richiama sovente la condizione di rischio associata all'affidamento consecutivo di contratti alle medesime aziende, che può trovare una giustificazione solo nel caso comporti un vantaggio economico competitivo. In generale, la condizione ipotetica di una stazione appaltante che aggiudichi i propri contratti ad aziende sempre diverse corrisponde ad una condizione di rischio nullo. Diversamente, quanto più alta è la proporzione di contratti aggiudicati da una stazione appaltante alla medesima azienda, maggiore è il rischio di corruzione.

Si consideri pertanto la distribuzione dei codici fiscali delle *K* aziende che sono risultate aggiudicatarie delle *N* procedure bandite da una certa stazione appaltante in un certo anno:

Azienda aggiudicataria	**Numero procedure aggiudicate**	**Quota di procedure aggiudicate**
1	n_1	f_1

2	n_2	f_2
…	…	…
k	n_k	f_k
…	…	…
K	n_K	f_K
	N	1

dove $f_k = n_k/N$. A livello di singola stazione appaltante, l'indicatore misura l'omogeneità di questa distribuzione, ottenuta invertendo l'indice di eterogeneità di Gini normalizzato (*E*).

L'omogeneità *O* della distribuzione si ottiene invertendo *E*, quindi $O = 1/E$. Infine, per ottenere il valore dell'indicatore per una determinata provincia, si procede con il calcolo della media della distribuzione degli *O* considerando le stazioni appaltanti localizzate in quella provincia.

14. Estensione del periodo di pubblicazione del bando (tra pubblicazione del bando e data di scadenza sottomissione proposte):

L'indicatore misura l'estensione temporale che intercorre tra la data di pubblicazione del bando e la data di scadenza per la presentazione delle offerte. La letteratura [121, 122,124,125,126] è infatti concorde nel ritenere che, da una parte, un periodo di pubblicità del bando estremamente ridotto possa rendere difficile (se non impossibile) la preparazione di offerte adeguate da parte delle aziende non collegate alla stazione appaltante da rapporti particolari, laddove invece le aziende collegate potrebbero venire informate in anticipo e informalmente dalla stazione appaltante. Dall'altra, un periodo di pubblicità del bando estremamente esteso potrebbe essere il segnale di incertezze sotto il profilo legale/procedurale che a loro volta possono segnalare un rischio di corruzione.
È uguale alla media aritmetica della differenza (in giorni) tra:

- data di scadenza di presentazione delle offerte;
- data di pubblicazione del bando.

15. Estensione del periodo di valutazione dell'offerta (tra data offerta e data aggiudicazione) per procedura di gara:

L'indicatore misura l'estensione temporale che intercorre tra la data di presentazione delle offerte da parte delle aziende e la data di aggiudicazione del contratto da parte della stazione appaltante. La letteratura [121,122,124,125,125] è concorde nel considerare a rischio di corruzione tanto intervalli di tempo per la valutazione delle offerte ridotti quanto estesi. Se, infatti, decisioni fulminee

potrebbero celare scelte premeditate e già prese, periodi di tempo prolungati potrebbero nascondere vere e proprie violazioni di legge e quindi un rischio di corruzione.
È calcolato come media aritmetica della differenza (in giorni) tra:

- data di aggiudicazione;
- data di scadenza di presentazione delle offerte.

16. Comportamenti "rischiosi" importi prossimi alla soglia v1

Rileva la frazione di contratti di importo compreso tra 37.500 e 40.000 € rispetto al numero di contratti di importo compreso tra 30.000 e 37.500 €. È uguale al rapporto tra:

- numeratore: numero di procedure con valore economico compreso tra 37.500 e 40.000 €;
- denominatore: numero di procedure con valore economico compreso tra 30.000 e 37.500 €.

L'indicatore ha la finalità di valutare se la stazione appaltante ha frazionato artificiosamente un determinato contratto con il solo scopo di non oltrepassare le soglie normativamente previste. Si tratta di una pratica nota nella letteratura scientifica internazionale con il termine di *contract splitting*, ovvero contratti multipli sotto soglia: è una pratica di frazionamento di progetti grandi (sotto il profilo del loro valore economico) in tanti contratti più piccoli, per evitare la competizione di aziende più grandi e livelli di controllo più stringenti e accurati previsti per bandi sopra soglie economiche stabilite (si vedano a titolo di esempio i documenti del World Bank Group: disponibili al link: https://www.worldbank.org/en/about/unit/integrity-vice-presidency).

17. Comportamenti "rischiosi" importi prossimi alla soglia v2

Rileva la frazione di contratti di importo compreso tra 20.000 e 40.000 € rispetto al numero di contratti di importo superiore a 40.000 €. È uguale al rapporto tra:

- numeratore: numero di procedure con valore economico compreso tra 20.000 e 40.000 €;
- denominatore: numero di procedure con valore economico superiore a 40.000 €.

Verso un indicatore composito di rischio di corruzione negli appalti pubblici?

Una delle principali sfide nella misurazione di fenomeni complessi e latenti, come corruzione e rischi di corruzione, consiste nel riassumere le informazioni disponibili espresse da un insieme di indicatori elementari, le cd. red flags, in un'unica metrica o Indicatore Composito (IC) di rischio di corruzione. Gli IC sono uno strumento di comunicazione utile per veicolare informazioni sintetiche in modo relativamente semplice. Sono ampiamente utilizzati in vari settori dei servizi pubblici come strumenti di analisi delle politiche e come veicolo di comunicazione pubblica per confrontare le unità di analisi (Paesi, regioni, aziende, ecc.). Tuttavia, nella costruzione di indicatori compositi, ci sono molte questioni metodologiche che necessitano di essere gestite con la massima cautela per evitare che i risultati presentati siano interpretati erroneamente e/o in maniera semplicistica. In particolare, è noto come lo sviluppo degli IC comporti una serie di fasi successive e complementari che presuppongono l'adozione e formulazione di scelte metodologiche, come ad esempio la scelta degli indicatori elementari da includere/escludere nel calcolo, la scelta dei criteri di

standardizzazione, la scelta degli eventuali pesi da applicare ai singoli indicatori, la scelta degli schemi di aggregazione, etc. Ciascuna di queste scelte riflette valutazioni sostanziali e altrettante fonti di incertezza dell'IC, incidendo sul valore numerico che assume e, in ultima istanza, sulla sua robustezza, validità e qualità.
Nel seguito, si sintetizzano le principali questioni di metodo che si impongono nel procedimento di calcolo di un indicatore composito.

Scelta del sistema di normalizzazione

La normalizzazione è necessaria prima dell'aggregazione dei dati quando i singoli indicatori elementari sono espressi in unità di misura diverse. La normalizzazione risponde in particolare a una duplice esigenza: i. esprimere nella stessa unità di misura gli indicatori red flags, in modo che assumano valori compresi nello stesso range e possano essere opportunamente confrontati e aggregati; ii. trasformare i valori originari degli indicatori rispetto ad una soglia di riferimento (ad esempio, la media) in modo tale che valori superiori alla soglia indichino diversi gradi di anomalia.
Esistono vari criteri per normalizzare indicatori semplici. I principali sono:

- Ranking, consistente nell'assegnazione di una posizione alle unità di analisi sulla base del valore assunto dall'indicatore composito;
- Standardizzazione (o z-score), che converte gli indicatori in una scala comune con una media di zero e una deviazione standard di uno;
- Min-Max, che normalizza gli indicatori in modo che risultino espressi nel medesimo intervallo [0, 1];
- Distanza da un valore di riferimento, che consiste nella misurazione della posizione relativa di un dato indicatore elementare rispetto al valore di riferimento;
- Scala categorica, che assegna un punteggio a ciascun indicatore. Le categorie possono essere numeriche – come quando i punteggi sono basati sui percentili della distribuzione dell'indicatore - oppure qualitative (ad es., "completamente raggiunto", "parzialmente raggiunto" o "non raggiunto").

Ciascuno dei precedenti criteri di normalizzazione presenta vantaggi e limiti in specifici contesti di analisi. Ad esempio, i metodi di normalizzazione che poggiano sulla definizione di soglie, portano con sé margini di arbitrarietà nella definizione del valore della soglia prescelta, malgrado il chiaro vantaggio della semplicità e anche se non sono influenzati da valori anomali. Il metodo della normalizzazione Min-Max, d'altra parte, presenta il limite di essere influenzato dalla presenza di valori estremi e/o anomali, che possono distorcere il valore assunto dall'indicatore trasformato.

Scelta degli schemi di ponderazione

Una questione molto controversa nel processo di costruzione di un IC è se e come assegnare pesi ai singoli indicatori elementari. Omettere di assegnare pesi alle varie componenti di un IC implicitamente comporta il riconoscimento di uguale importanza assegnata a ciascuna di esse. Diversamente, si può assegnare uno schema di ponderazione diverso per dare maggiore importanza alle componenti considerate più rilevanti.
I pesi possono provenire da diverse fonti, per esempio da opinioni personali di ricercatori, esperti o società civile, anche ricavate attraverso sondaggi. In alternativa, i pesi possono essere determinati attraverso il ricorso a tecniche statistiche. Pertanto, parallelamente agli approcci normativi, che dipendono da giudizi di valore soggettivamente determinati, altri criteri propongono l'assegnazione

di pesi in modo endogeno, ovvero a partire dai dati e in funzione di essi, attraverso tecniche statistiche. Le principali tecniche statistiche al riguardo sono Analisi delle Componenti Principali (PCA), Analisi Fattoriale e Benefit-of-the-Doubt. Parte della letteratura sull'argomento è incline a preferire schemi di ponderazione non soggettivi e sostiene che il compito di assegnare pesi dovrebbe essere analiticamente valido, trasparente e coerente con l'obiettivo della ricerca.

Scelta del sistema di aggregazione

Esistono diversi criteri per aggregare i singoli indicatori elementari in un IC. L'approccio compensativo si applica generalmente quando gli indicatori elementari sono considerati sostituibili, talché un deficit in un indicatore può essere compensato da un'eccedenza in un altro. In questo caso vengono adottate funzioni lineari, come la media aritmetica. L'approccio non compensativo si applica invece quando gli indicatori elementari sono considerati non sostituibili. In questo caso vengono adottate funzioni non lineari che tengono conto - implicitamente o esplicitamente - dello squilibrio tra valori diversi in termini di penalizzazione. Al riguardo, in letteratura (si veda per esempio OCSE 2008a, b) si osserva come il metodo di aggregazione lineare sia opportunamente impiegabile quando i singoli indicatori hanno la stessa unità di misura, mentre le aggregazioni di tipo geometrico sono più adeguate quando è richiesto un certo grado di non compensabilità tra singoli indicatori. In generale, si distinguono due tipologie di metodi di aggregazione: l'approccio compensativo e l'approccio non compensativo.
Nella costruzione di un indicatore sintetico di rischio di corruzione è necessario prestare attenzione a questi e altri aspetti specifici/tecnici che hanno conseguenze sostanziali nella scelta delle modalità di aggregazione, tra cui (ma non solo): i. l'ipotesi generale di una relazione positiva degli indicatori semplici con il rischio di corruzione può trovare eccezioni in circostanze specifiche; ii. l'ipotesi generale di una relazione lineare tra indicatori semplici e rischio di corruzione potrebbe non essere verificata in situazioni specifiche, per esempio quando il rischio di corruzione mostra "salti" attorno a determinate soglie economiche.
L'attuale letteratura che propone IC di rischio di corruzione negli appalti pubblici [124] ha scelto di fare affidamento su metodi di aggregazione lineare, sommando o calcolando la media semplice dei singoli indicatori red flag per costruire la misura sintetica di rischio di corruzione in appalti pubblici.

Analisi delle relazioni tra indicatori elementari e della struttura di dimensionalità dei dati

La corruzione è un fenomeno complesso e latente, cioè non direttamente osservabile nei dati, data la sua natura sommersa. Nonostante la sua complessità, tanto la corruzione quanto il rischio di corruzione sono sovente trattati e misurati come costrutti unidimensionali. Infatti, l'assunto implicito di qualsiasi tentativo di misurazione volto a sviluppare un indicatore composito di rischio di corruzione è che le singole red flag siano ciascuna espressione dello stesso unico fenomeno sottostante (corruzione). Tuttavia, è ragionevole pensare che diverse sottodimensioni (o sottogruppi) di red flag misurino, da punti di vista diversi, diverse sfaccettature dello stesso fenomeno sottostante. Se questo è il caso, i singoli indicatori red flags tenderebbero ad essere altamente correlati all'interno dei sottogruppi e non correlati tra sottogruppi.
Riconoscere la complessità (o multidimensionalità) di fenomeni latenti quali il rischio di corruzione porta con sé una serie di scelte metodologiche. Una prima questione centrale riguarda l'individuazione delle sottodimensioni di cui tener conto. Esistono diversi strumenti metodologico-statistici per selezionare numero e composizione interna delle dimensioni in cui possono essere raggruppati i singoli indicatori elementari. Pertanto, la questione chiave diventa la scelta del metodo più appropriato per inferire la struttura di dimensionalità e relazionale degli indicatori red flags. Le

tecniche più diffuse sono l'analisi delle componenti principali e l'analisi fattoriale. Entrambe sono tecniche di riduzione della complessità dei dati, ma mentre il primo consente l'estrazione di misure composite come combinazione lineare delle variabili osservate, il secondo si basa su un modello statistico formale che consente di stimare le variabili osservate da fattori latenti. Inoltre, i metodi di riduzione della complessità dei dati relativi alla struttura delle variabili possono essere esplorativi e confermativi. Questi ultimi individuano ex-ante il numero e la composizione interna delle dimensioni (ad esempio, specificando quali indicatori elementari ne fanno parte) e attraverso la tecnica prescelta sottopongono a verifica di plausibilità la struttura di dimensionalità ipotizzata. Pertanto, per applicarli, è necessario specificare in anticipo una struttura dimensionale. Tale indicazione può essere difficile da trovare nella pratica; ciò accade ogniqualvolta non si disponga di informazioni preliminari sulla struttura dimensionale dei dati in questione. In questi ultimi casi possono essere adottate le tecniche esplorative, che consentono di esplorare la struttura dimensionale senza alcuna ipotesi di partenza.
Infine, nell'ambito della valutazione del rischio di corruzione, i modelli statistici che assumono che le associazioni tra singoli indicatori o red flags siano rappresentate da un costrutto latente multidimensionale (es. corruzione) possono tutti essere considerati adeguati allo scopo di studiarne la struttura relazionale. In particolare, i modelli a variabili latenti – come i modelli Structural Equation Modeling (SEM) e Item Response Theory (IRT) – sono tra i più promettenti in quanto consentono di studiare la struttura relazionale degli indicatori semplici osservati spiegando l'aspetto latente e multidimensionale del rischio di corruzione.

Tecniche di analisi di dati proveniente da fonti aperte basate su altri algoritmi

Nel contesto di OSINT, le tecniche della teoria dei grafi possono essere utilizzate per rappresentare e analizzare le relazioni tra individui, gruppi ed entità di interesse.
Uno degli usi più comuni della teoria dei grafi in OSINT è la creazione di grafici di social network. I grafici dei social network vengono utilizzati per rappresentare le connessioni tra individui o gruppi in un social network.
In un contesto OSINT, i grafi dei social network possono essere utilizzati per mappare le relazioni tra le persone coinvolte in un particolare evento o attività, come una rete criminale o un'organizzazione terroristica. Questi grafici possono rivelare informazioni importanti sulla struttura e le dinamiche della rete, inclusi gli attori chiave e le loro relazioni, nonché i modelli di comunicazione e coordinamento tra di loro.

Un'altra applicazione della teoria dei grafi in OSINT è nell'analisi dei collegamenti. L'analisi dei collegamenti comporta l'analisi delle relazioni tra entità di interesse in base ai collegamenti o alle connessioni tra di esse. In un contesto OSINT, l'analisi dei collegamenti può essere utilizzata per identificare modelli di attività o comportamento, come il riciclaggio di denaro o il finanziamento del terrorismo [131].
Mappando le relazioni tra entità e analizzando il flusso di fondi o informazioni tra di loro, gli investigatori possono ottenere informazioni sulla struttura e l'organizzazione della rete e identificare attori o nodi chiave.

Inoltre le tecniche possono essere combinate per creare analisi più avanzate e sofisticate. Un esempio di ciò è la creazione di una metrica di somiglianza univoca che soppesa le relazioni del grafico in base a molteplici aspetti delle relazioni.
Ad esempio, nel contesto di OSINT, è possibile definire una metrica di somiglianza multidimensionale per identificare campagne dannose su piattaforme di social media come Twitter.

La metrica di somiglianza può essere misurata come una combinazione lineare della somiglianza temporale dei tweet inviati e del punteggio di attribuzione dell'autore relativo ai profili dei mittenti. Come spiegato da [132], la somiglianza temporale dei tweet inviati può essere utilizzata per identificare modelli di attività e comportamento, mentre il punteggio di attribuzione dell'autore può essere utilizzato per identificare l'autenticità dei profili e rilevare eventuali attività dannose o tentativi di impersonificazione.
Combinando questi diversi aspetti delle relazioni nel grafico, gli investigatori possono ottenere una comprensione più completa e accurata della rete e identificare potenziali minacce o rischi in modo più efficace. Questo approccio evidenzia anche l'importanza di considerare più dimensioni delle relazioni in un grafico per condurre un'analisi OSINT efficace.

Ci sono una serie di statistiche e misure associate all'analisi della rete che forniscono informazioni che ci aiutano a capire come le posizioni delle diverse entità nella rete influenzano il suo funzionamento [133,134], per esempio, ci sono diverse misure di centralità che possono essere utili per l'analisi OSINT.
La centralità dell'autovettore, ad esempio, presuppone che avere contatti importanti sia più critico che avere un gran numero di contatti. Sotto questa ipotesi, la centralità di un nodo è correlata alla somma della centralità dei suoi vicini, e le entità con un'elevata centralità di Autovettore sono quelle che sono connesse ad altre entità importanti nella rete [135,136].
PageRank è un'altra misura di centralità originariamente sviluppata da Google per classificare le pagine Web nei risultati di ricerca. Utilizza un approccio simile alla centralità degli autovettori, ma tiene anche conto della qualità e dell'importanza delle pagine che si collegano a una particolare pagina [137,138].

Inoltre, gli algoritmi di rilevamento della comunità possono essere utilizzati per identificare cluster o gruppi di nodi che sono maggiormente collegati tra loro rispetto al resto della rete. Questi algoritmi possono fornire informazioni sulla struttura della rete e identificare potenziali sottogruppi o comunità che possono avere proprietà o comportamenti unici [139, 140].
Nel complesso, l'uso di misure di centralità e algoritmi di rilevamento della comunità in OSINT può aiutare a identificare i principali attori e sottogruppi all'interno di una rete e a capire come le informazioni fluiscono attraverso la rete. Combinando queste tecniche con altri metodi OSINT, gli investigatori possono acquisire una comprensione più completa della rete e dei suoi potenziali rischi e minacce.

Bibliografia

[1] A. J. P, 'Computer security threat monitoring and surveillance', *Technical Report, James P. Anderson Company*, 1980, Accessed: Jan. 29, 2023. [Online]. Available: https://cir.nii.ac.jp/crid/1573950399661362176

[2] M. C. Huebscher and J. A. McCann, 'A survey of autonomic computing—degrees, models, and applications', *ACM Comput. Surv.*, vol. 40, no. 3, p. 7:1-7:28, Aug. 2008, doi: 10.1145/1380584.1380585.

[3] H. Debar, 'Section: Security Operations & Incident Management', in *The Cyber Security Body of Knowledge v1.1.0, 2021*, University of Bristol, 2021. [Online]. Available: https://www.cybok.org/

[4] R. Lippmann, J. W. Haines, D. J. Fried, J. Korba, and K. Das, 'The 1999 DARPA off-line intrusion detection evaluation', *Computer Networks*, vol. 34, no. 4, pp. 579–595, Oct. 2000, doi: 10.1016/S1389-1286(00)00139-0.

[5] J. McHugh, 'Testing Intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory', *ACM Trans. Inf. Syst. Secur.*, vol. 3, no. 4, pp. 262–294, Nov. 2000, doi: 10.1145/382912.382923.

[6] M. Wood and M. Erlinger, 'Intrusion Detection Message Exchange Requirements', RFC Editor, RFC 4766, Mar. 2007. [Online]. Available: https://www.rfc-editor.org/info/rfc4766

[7] H. Debar, D. Curry, and B. Feinstein, 'The Intrusion Detection Message Exchange Format (IDMEF)', RFC Editor, RFC 4765, Mar. 2007. [Online]. Available: http://www.rfc-editor.org/rfc/rfc4765.txt

[8] B. Feinstein and G. Matthews, 'The Intrusion Detection Exchange Protocol (IDXP)', RFC Editor, RFC 4767, Mar. 2007. [Online]. Available: https://www.rfc-editor.org/info/rfc4767

[9] R. Gerhards, 'The Syslog Protocol', RFC Editor, RFC 5424, Mar. 2009. [Online]. Available: http://www.rfc-editor.org/rfc/rfc5424.txt

[10] P. Radoglou-Grammatikis *et al.*, 'SPEAR SIEM: A Security Information and Event Management system for the Smart Grid', *Computer Networks*, vol. 193, p. 108008, Jul. 2021, doi: 10.1016/j.comnet.2021.108008.

[11] G. González-Granadillo, S. González-Zarzosa, and R. Diaz, 'Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures', *Sensors*, vol. 21, no. 14, Art. no. 14, Jan. 2021, doi: 10.3390/s21144759.

[12] M. Cinque, D. Cotroneo, and A. Pecchia, 'Challenges and Directions in Security Information and Event Management (SIEM)', in *2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*, Oct. 2018, pp. 95–99. doi: 10.1109/ISSREW.2018.00-24.

[13] H. Debar and A. Wespi, 'Aggregation and Correlation of Intrusion-Detection Alerts', in *Recent Advances in Intrusion Detection*, Berlin, Heidelberg, 2001, pp. 85–103. doi: 10.1007/3-540-45474-8_6.

[14] F. Cuppens and A. Miege, 'Alert correlation in a cooperative intrusion detection framework', in *Proceedings 2002 IEEE Symposium on Security and Privacy*, May 2002, pp. 202–215. doi: 10.1109/SECPRI.2002.1004372.

[15] P. Ning, Y. Cui, and D. S. Reeves, 'Constructing attack scenarios through correlation of intrusion alerts', in *Proceedings of the 9th ACM conference on Computer and communications security*, New York, NY, USA, Nov. 2002, pp. 245–254. doi: 10.1145/586110.586144.

[16] J. Zhou, M. Heckman, B. Reynolds, A. Carlson, and M. Bishop, 'Modeling network intrusion detection alerts for correlation', *ACM Trans. Inf. Syst. Secur.*, vol. 10, no. 1, pp. 4-es, Feb. 2007, doi: 10.1145/1210263.1210267.

[17] A. Valdes and K. Skinner, 'Probabilistic Alert Correlation', in *Proceedings of the 4th International Symposium on Recent Advances in Intrusion Detection*, Berlin, Heidelberg, Oct. 2001, pp. 54–68.

[18] X. Yin, W. Yurcik, M. Treaster, Y. Li, and K. Lakkaraju, 'VisFlowConnect: netflow visualizations of link relationships for security situational awareness', in *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security*, New York, NY, USA, Oct. 2004, pp. 26–34. doi: 10.1145/1029208.1029214.

[19] K. Julisch and M. Dacier, 'Mining intrusion detection alarms for actionable knowledge', in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA, Jul. 2002, pp. 366–375. doi: 10.1145/775047.775101

[20] I. Kotenko and A. Chechulin, 'Attack modeling and security evaluation in SIEM systems', *International Transactions on Systems Science and Applications*, vol. 8, pp. 129–147, 2012.

[21] O. Wenge, U. Lampe, C. Rensing, and R. Steinmetz, 'Security Information and Event Monitoring as a Service: a Survey on Current Concerns and Solutions', *PIK - Praxis der Informationsverarbeitung und Kommunikation*, vol. 37, no. 2, pp. 163–170, Jun. 2014, doi: 10.1515/pik-2014-0009.

[22] J. M. López Velásquez, S. M. Martínez Monterrubio, L. E. Sánchez Crespo, and D. Garcia Rosado, 'Systematic review of SIEM technology: SIEM-SC birth', *Int. J. Inf. Secur.*, Jan. 2023, doi: 10.1007/s10207-022-00657-9.

[23] 'A GDPR Compliant SIEM Solution - ProQuest'. https://www.proquest.com/openview/caf2d8e6b796bc5ab720f4534393ba0a/1?cbl=396497&parentSessionId=SadEgQHLoUD6hPZAVQqY0cIjbYbJD1UfrkNmfD53l%2F4%3D&pq-origsite=gscholar&parentSessionId=eVvuyAoj5SGWi3lrt7tydfL6VgqWAWosQhXR5U6dR48%3D (accessed Mar. 06, 2023).

[24] I. Kotenko, D. Gaifulina, and I. Zelichenok, 'Systematic Literature Review of Security Event Correlation Methods', *IEEE Access*, vol. 10, pp. 43387–43420, 2022, doi: 10.1109/ACCESS.2022.3168976.

[25] D. Levshun and I. Kotenko, 'A survey on artificial intelligence techniques for security event correlation: models, challenges, and opportunities', *Artif Intell Rev*, Jan. 2023, doi: 10.1007/s10462-022-10381-4.

[26] H. J. Williams and I. Blum, 'Defining Second Generation Open Source Intelligence (OSINT) for the Defense Enterprise', *Research Reports RAND Corporation*, Santa Monica, CA, 2018. [Online]. Available: https://www.rand.org/pubs/research_reports/RR1964.html

[27] D. M. Bornn, 'Service members, civilians learn to harness power of 'Open Source' information', U.S. Army www.army.mil, 9 January 2013. [Online]. Available https://www.army.mil/article/94007/Service_members__civilians_learn_to_harness_power_of__Open_Source__information

[28] Publications Office of the European Union, 'Open-source intelligence', data.europa.eu - The official portal for European data, 2 May 2022. [Online]. Available: https://data.europa.eu/en/publications/datastories/open-source-intelligence

[29] osintforukraine.com, 'OSINT for Ukraine', OSINT for Ukraine, Accessed on 20 March 2023. [Online]. Available: https://www.osintforukraine.com/

[30] R. A. Best and A. Cumming, 'Open Source Intelligence (OSINT): Issues for Congress', Congressional Research Service (CRS), Report for Congress, 5 December 2007. [Online]. Available https://sgp.fas.org/crs/intel/RL34270.pdf

[31] R. A. Best and A. Cumming, 'Open Source Intelligence (OSINT): Issues for Congress', Congressional Research Service (CRS), Report for Congress, 28 January 2008. [Online]. Available https://web.archive.org/web/20160304031047/http://www.osint.org/crs-report-osint.pdf
[32] M. M. Lowenthal, 'Open-Source Intelligence: New Myths, New Realities', in R. Z. George and R. D. Kline (eds.), Intelligence and the national security strategist: enduring issues and challenges, Lanham: Rowman and Littlefield, ISBN-13 9780742540392, 2005
[33] N. Asciutti, A. Fraticelli, E. Iommi, 'Il mondo dell'Osint', technical report, Facoltà di Scienze e Tecnologie, Università degli Studi di Camerino, 2022. [Online]. Available https://computerscience.unicam.it/marcantoni/project/Il mondo OSINT.pdf
[34] Public Law 109-163, 'National Defense Authorization Act for Fiscal Year 2006, Sec. 931, Department of Defense Strategy for Open-Source Intelligence', 6 January 2006. [Online]. Available: https://www.govinfo.gov/app/details/PLAW-109publ163/
[35] nso.nato.int, 'NATOTerm The Official NATO Terminology Database', NATOTermOTAN, Approval date 31 October 2013, Accessed on 20 March 2023. [Online]. Available: https://nso.nato.int/natoterm/content/nato/pages/home.html
[36] J. T. Richelson, 'The U.S. Intelligence Community', 7th ed. Westview Press, Boulder, CO, USA, ISBN-13: 9780813349190, 2016
[37] 'Joint Publication 2-0, Joint Intelligence', Defense Technical Information Center (DTIC), Department of Defense, 2013. [Online]. Available: https://irp.fas.org/doddir/dod/jp2_0.pdf
[38] L. K. Johnson editor. 'Handbook of Intelligence Studies', 1st ed. Routledge, New York, NY, USA, ISBN-13: 978-0-415-77050-7, 2007
[39] H. Gibson, 'Acquisition and preparation of data for osint investigations'. In B. Akhgar, P. S. Bayerl, and F. Sampson (eds.), *Open source intelligence investigation: From Strategy to Implementation*, pp 69–93. Springer, Basilea, Svizzera, 2016.
[40] B. Akhgar, P. S. Bayerl, and F. Sampson, editors. 'Open Source Intelligence Investigation: From Strategy to Implementation', *Advanced Sciences and Technologies for Security Applications*, 1st ed. Springer, Cham, Switzerland, ISBN-13: 978-3-319-47671-1, 2017, doi: 10.1007/978-3-319-47671-1
[41] Central Intelligence Agency (CIA), 'The Intelligence Cycle', 23 March 2013. [Online]. Available: https://web.archive.org/web/20200508151219/https://www.cia.gov/kids-page/6-12th-grade/who-we-are-what-we-do/the-intelligence-cycle.html
[42] M. Bazzell, 'OSINT Techniques: Resources for Uncovering Online Information', 10th ed. Independently published, ISBN-13: 979-8366360401, 2023
[43] M. Bazzell, 'Open source intelligence techniques: resources for searching and analyzing Online information', 6th ed. Independently published, ISBN-13: 978-1984201577, 2018
[44] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. Christiano, J. Leike, R. Lowe, 'Training language models to follow instructions with human feedback', 2022. [Online] Available https://arxiv.org/abs/2203.02155
[45] L. Benes, 'OSINT, New Technologies, Education: Expanding Opportunities and Threats. A New Paradigm', *Journal of Strategic Security*, 6, no. 3 Suppl. (2013): 22-37. [Online]. Available: http://dx.doi.org/10.5038/1944-0472.6.3S.3
[46] Levshun, D., Kotenko, I. (2023). A survey on artificial intelligence techniques for security event correlation: models, challenges, and opportunities. Artificial Intelligence Review:1-44
[47] Hostiadi D. P., Susila M. D., Huizen R. R. (2019) A new alert correlation model based on similarity approach. In: 2019 1st International Conference on Cybernetics and Intelligent System (ICORIS), IEEE, pp 133–137

[48] Sun J, Gu L, Chen K (2020) An efficient alert aggregation method based on conditional rough entropy and knowledge granularity. Entropy 22(3):324
[49] Kotenko I, Fedorchenko A, Saenko I, et al Parallelization of security event correlation based on accounting of event type links. In: 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP), IEEE, pp 462–469
[50] Kotenko I, Fedorchenko A, Doynikova E (2020) Data analytics for security management of complex heterogeneous systems: Event correlation and security assessment tasks. In: Advances in Cyber Security Analytics and Decision Systems. Springer, Cham, pp 79–116
[51] Mahdavi E, Fanian A, Amini F (2020) A real-time alert correlation method based on code-books for intrusion detection systems. Computers & Security 89:101,661
[52] Siddiqui AJ, Boukerche A (2021) TempoCode-IoT: temporal codebook-based encoding of flow features for intrusion detection in internet of things. Cluster Computing 24(1):17–35
[53] Tao Xl, Shi L, Zhao F, et al (2021) A hybrid alarm association method based on AP clustering and causality. Wireless Communications and Mobile Computing 2021(5):1–10
[54] Lanoe D, Hurfin M, Totel E (2018) A scalable and efficient correlation engine to detect multi-step attacks in distributed systems. In: 2018 IEEE 37th Symposium on Reliable Distributed Systems (SRDS), IEEE, pp 31–40
[55] Bénard C, Biau G, Da Veiga S, et al (2021) SIRIUS: Stable and interpretable rule set for classification. Electronic Journal of Statistics 15(1):427–505
[56] Dhaou A, Bertoncello A, Gourvénec S, et al (2021) Causal and interpretable rules for time series analysis. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, pp 2764–2772
[57] Xie T, Zheng Q, Zhang W (2018) Mining temporal characteristics of behaviors from interval events in e-learning. Information Sciences 447:169–185
[58] Eckmann S. T., Vigna G, Kemmerer R. A. (2002) Statl: An attack language for state-based intrusion detection. Journal of computer security 10(1-2):71–103
[59] Tidjon L. N., Frappier M, Mammar A (2020) Intrusion detection using ASTDs. In: International Conference on Advanced Information Networking and Applications. Springer, Cham, pp 1397–1411
[60] Meier M, Bischof N, Holz T (2002) SHEDEL - a simple hierarchical event description language for specifying attack signatures. In: Security in the Information Society. Springer US, Boston, MA, p 559–571
[61] Jaeger D, Ussath M, Cheng F, et al (2015) Multi-step attack pattern detection on normalized event logs. In: 2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing, IEEE, pp 390–398
[62] Almseidin M, Piller I, Al-Kasassbeh M, et al (2019) Fuzzy automaton as a detection mechanism for the multi-step attack. International Journal on Advanced Science, Engineering and Information Technology 9(2):575–586
[63] Mikolov T, Chen K, Corrado G, et al (2013) Efficient estimation of word representations in vector space. In: 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, pp 1–12
[64] Mikolov T, Sutskever I, Chen K, et al (2013) Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems 26:1–9
[65] Pennington J, Socher R, Manning CD (2014) Glove: Global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
[66] Wang J, Tang Y, He S, et al (2020) LogEvent2vec: Logevent-to-vector based anomaly detection for large-scale logs in internet of things. Sensors 20(9):2451

[67] Wang J, Zhao C, He S, et al (2022) LogUAD: Log unsupervised anomaly detection based on Word2Vec. Computer Systems Science and Engineering 41(3):1207–1222
[68] Le Q, Mikolov T (2014) Distributed representations of sentences and documents. In: International conference on machine learning, PMLR, pp 1188–1196
[69] Liu L, Chen C, Zhang J Doc2vec-based insider threat detection through behaviour analysis of multi-source security logs. In: 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), IEEE, pp 301–309
[70] Chen R, Zhang S, Li DLogtransfer: Cross-system log anomaly detection for software systems with transfer learning. In: 2020 IEEE 31st International Symposium on Software Reliability Engineering (ISSRE), IEEE, pp 37–47
[71] Peters ME, Neumann M, Iyyer M, et al (2018) Deep contextualized word representations. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. ACL, New Orleans, Louisiana, pp 2227–2237
[72] Devlin J, Chang MW, Lee K, et al (2019) BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), pp 4171–4186
[73] Zhan Y, Haddadi H (2019) Towards automating smart homes: Contextual and temporal dynamics of activity prediction. In: Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, pp 413–417
[74] Seyyar YE, Yavuz AG, Unver HM (2022) An attack detection framework based on BERT and deep learning. IEEE Access Early Access:1–13
[75] Huang L, Ji H, Cho K, et al (2018) Zero-shot transfer learning for event extrac- tion. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, pp 2160–2170
[76] Deng S, Zhang N, Li L, et al (2021) OntoED: Low-resource event detection with ontology embedding. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics, pp 2828–2839
[77] Barzegar M, Shajari M (2018) Attack scenario reconstruction using intrusion semantics. Expert Systems with Applications 108:119–133
[78] Zheng H, Wang Y, Han C, et al (2018) Learning and applying ontology for machine learning in cyber attack detection. In: 2018 17th IEEE Inter- national Conference on Trust, Security and Privacy in Computing and Communications, IEEE, pp 1309–1315
[79] Wang Q, Jiang J, Shi Z, et al (2018) A novel multi-source fusion model for known and unknown attack scenarios. In: 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), IEEE, pp 727–736
[80] Hassan WU, Guo S, Li D, et al (2019) Nodoze: Combatting threat alert fatigue with automated provenance triage. In: Network and Distributed Systems Security Symposium, pp 1–15
[81] Hassan WU, Noureddine MA, Datta P, et al (2020) OmegaLog: High-fidelity attack investigation via transparent multi-layer log analysis. In: Network and Distributed System Security Symposium, pp 1–16
[82] Han X, Pasquier T, Bates A, et al (2020) UNICORN: Runtime provenance- based detector for advanced persistent threats. In: Network and Distributed System Security Symposium, pp 1–18

[83] Milajerdi SM, Gjomemo R, Eshete B, et al (2019) Holmes: real-time apt detection through correlation of suspicious information flows. In: 2019 IEEE Symposium on Security and Privacy (SP), IEEE, pp 1137–1152
[84] Zeng J, Chua ZL, Chen Y, et al (2021) Watson: Abstracting behaviors from audit logs via aggregation of contextual semantics. In: Proceedings of the 28th Annual Network and Distributed System Security Symposium, NDSS, pp 1–18
[85] Nicart, Esther, et al. "Building document treatment chains using reinforcement learning and intuitive feedback." 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI). IEEE, 2016.
[86] Yang, Hsin-Chang, and Chung-Hong Lee. "Mining open source text documents for intelligence gathering." 2012 International Symposium on Information Technologies in Medicine and Education. Vol. 2. IEEE, 2012.
[87] Noubours, Sandra, Albert Pritzkau, and Ulrich Schade. "NLP as an essential ingredient of effective OSINT frameworks." 2013 Military Communications and Information Systems Conference. IEEE, 2013.
[88] J. Evangelista, et al. (2020): Systematic Literature Review to Investigate the Application of Open Source Intelligence (OSINT) with Artificial Intelligence, Journal of Applied Security Research, DOI: 10.1080/19361610.2020.1761737
[89] R.Layton, and P.A. Watters. Automating open source intelligence: Algorithms for OSINT. Syngress, 2015.
[90] Chaudhary, Megha, and Divya Bansal. "Open source intelligence extraction for terrorism-related information: A review." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 12.5 (2022): e1473.
[91] Riebe, Thea, et al. "Cysecalert: An alert generation system for cyber security events using open source intelligence data." Information and Communications Security: 23rd International Conference, ICICS 2021.
[92] Bayer, Markus, et al. "CySecBERT: A Domain-Adapted Language Model for the Cybersecurity Domain." arXiv preprint arXiv:2212.02974 (2022).
[93] Chang YC, Wang SD (2016) The concept of attack scenarios and its applications in Android malware detection. In: 2016 IEEE 18th International Conference on High Performance Computing and Communications, IEEE, pp 1485–1492
[94] Kotenko I, Saenko I, Branitskiy A (2018b) Framework for mobile Internet of things security monitoring based on Big Data processing and machine learning. IEEE Access 6:72,714–72,723
[95] Li S, Zhang Q, Wu X, et al (2021) Attribution classification method of APT malware in iot using machine learning techniques. Security and Communication Networks 2021:1–12
[96] Shen Y, Mariconti E, Vervier PA, et al (2018) Tiresias: Predicting security events through deep learning. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, pp 592–605
[97] Du M, Li F, Zheng G, et al (2017) Deeplog: Anomaly detection and diagnosis from system logs through deep learning. In: Proceedings of the 2017 ACM SIGSAC conference on computer and communications security, pp 1285– 1298
[98] Wang Z, Chen Z, Ni J, et al (2021b) Multi-scale one-class recurrent neural networks for discrete event sequence anomaly detection. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, pp 3726–3734
[99] Chen H, Xiao R, Jin S (2020a) Real-time detection of cloud tenant malicious behavior based on CNN. In: 2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom), IEEE, pp 998–1005

[100] Abdullayeva FJ (2021) Advanced persistent threat attack detection method in cloud computing based on autoencoder and softmax regression algorithm. Array 10:100,067
[101] Min B, Yoo J, Kim S, et al (2021) Network anomaly detection using memory- augmented deep autoencoder. IEEE Access 9:104,695–104,706
[102] Liu J, Chen Y, Liu K (2019b) Exploiting the ground-truth: An adversarial imitation based knowledge distillation approach for event detection. Pro- ceedings of the AAAI Conference on Artificial Intelligence 33(01):6754–6761
[103] Oki M, Takeuchi K, Uematsu Y (2018) Mobile network failure event detection and forecasting with multiple user activity data sets. Proceedings of the AAAI Conference on Artificial Intelligence 32(1):7786–7792
[104] Ghafouri A, Vorobeychik Y, Koutsoukos X (2018) Adversarial regression for detecting attacks in cyber-physical systems. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence. AAAI Press, Stockholm, Sweden, IJCAI'18, pp 3769—-3775
[105] Joloudari JH, Haderbadi M, Mashmool A, et al (2020) Early detection of the advanced persistent threat attack using performance analysis of deep learning. IEEE Access 8:186,125–186,137
[106] Cheng H, Xie Z, Shi Y, et al (2019) Multi-step data prediction in wireless sensor networks based on one-dimensional CNN and bidirectional LSTM. IEEE Access 7:117,883–117,896
[107] Do Xuan C, Dao MH (2021) A novel approach for APT attack detection based on combined deep learning model. Neural Computing and Applications 33(20):13,251–13,264
[108] Haas S, Fischer M (2019) On the alert correlation process for the detection of multi-step attacks and a graph-based realization. ACM SIGAPP Applied Computing Review 19(1):5–19
[109] Bajtoš T, Sokol P, Mézešová T (2020) Multi-stage cyber-attacks detection in the industrial control systems. In: Recent Developments on Industrial Control Systems Resilience. Springer, Cham, p 151–173
[110] Wang X, Gong X, Yu L, et al (2021a) MAAC: Novel alert correlation method to detect multi-step attack. In: 2021 IEEE 20th International Confer- ence on Trust, Security and Privacy in Computing and Communications (TrustCom), IEEE, pp 726–733
[111] Zhang X, Wu T, Zheng Q, et al (2022) Multi-step attack detection based on pre-trained hidden Markov models. Sensors 22(8):2874
[112] Liu F, Wen Y, Zhang D, et al (2019a) Log2vec: A heterogeneous graph embedding based approach for detecting cyber threats within enterprise. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, pp 1777–1794
[113] Mao B, Liu J, Lai Y, et al (2021) Mif: A multi-step attack scenario recon- struction and attack chains extraction method based on multi-information fusion. Computer Networks 198:108,340
[114] Luo W, Zhang H, Yang X, et al (2020) Dynamic heterogeneous graph neural network for real-time event prediction. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, p 3213–3223
[115] Nguyen T, Grishman R (2018) Graph convolutional networks with argument- aware pooling for event detection. Proceedings of the AAAI Conference on Artificial Intelligence 32(1):5900–5907
[116] Deng A, Hooi B (2021) Graph neural network-based anomaly detection in multivariate time series. Proceedings of the AAAI Conference on Artificial Intelligence 35(5):4027–4035
[117] Analisi istruttoria per l'individuazione di indicatori di rischio corruzione e di prevenzione e contrasto nelle amministrazioni pubbliche. Report finale ANAC, 2017
[118] Decarolis, Francesco, Ray Fisman, Paolo Pinotti, and Silvia Vannutelli (2019). Corruption in Procurement: New facts from Italian Government Contracting. Working Paper.

[119] Abdou, A., Ágnes Czibik, A., Tóth, B. and Fazekas, M. (2021). COVID-19 emergency public procurement in Romania: Corruption risks and market behavior. Working Paper series: GTI-WP/2021:03. Budapest, Hungary.

[120] Auriol, Emmanuelle, StÅLephane Straub, and Thomas Flochel. 2016. “Public procurement and rent-seeking: the case of Paraguay.” World Development 77:395–407.

[121] Fazekas, M., Cingolani, L., & Tóth, B. (2017). A comprehensive review of objective corruption proxies in public procurement: risky actors, transactions, and vehicles of rent extraction. Government Transparency Institute Working Paper Series No. GTI-WP/2016:03, Budapest.

[122] European Anti-Fraud Office - OLAF (2017). Fraud in Public Procurement. A collection of Red Flags and Best Practices. Directorate D: Policy Unit D.2 Fraud Prevention, Reporting and Analysis.

[123] Fazekas, Mihaly, Sberna, Salvatore, Vannucci, Alberto (2021). The extra-legal governance of corruption: Tracing the organization of corruption in public procurement. Governance.

[124] Fazekas, M., Tóth, I. J., & King, L. P. (2016). An Objective Corruption Risk Index Using Public Procurement Data. European Journal of Criminal Policy and Research, 22(3), 369–397.

[125] European Anti-Fraud Office (OLAF). Procurement: costs we pay for corruption. Identifying and Reducing Corruption in Public Procurement in the EU. European Anti-Fraud Office (OLAF), PwC EU Services and Ecorys, and University of Utrecht. 2013

[126] Fazekas, M., & Kocsis, G. (2015). Uncovering High-Level Corruption: Cross-National Corruption Proxies Using Government Contracting Data. GTI-WP/2015:02, Budapest: Government Transparency Institute.

[127] Klasnja, Marko. 2015. “Corruption and the incumbency disadvantage: theory and evidence.” The Journal of Politics 77 (4): 928–942.

[128] Ferwerda, Deleanu, Unger (2017). Corruption in Public Procurement: Finding the Right Indicators. Eur J Crim Policy Res, 23:245–267.

[129] Ferraris, Mazza and Scomparin (Eds.). Warning on Crime (WoC). Preventing and Combatting Crime in Public Procurement (www.warningoncrime.eu) Torino: Università degli Studi di Torino, 2016.

[130] Scomparin, L. (Eds.). Corruzione e infiltrazioni criminali negli appalti pubblici. Strumenti di prevenzione e contrasto. Giappichelli Editore, 2016.

[131] Chen, Xucan, et al. "Identifying Darknet Vendor Wallets by Matching Feedback Reviews with Bitcoin Transactions." 2021 International Conference on Data Mining Workshops (ICDMW). IEEE, 2021

[132] Perez, Charles, et al. "REPLOT: retrieving profile links on twitter for suspicious networks detection." Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2013.

[133] Krebs, Valdis. "Social network analysis: an introduction." In: OrgNet. 2013

[134] Chaudhary, Megha, and Divya Bansal. "Open source intelligence extraction for terrorism-related information: A review." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 12.5 (2022): e1473.

[135] Berzinji, Ala, Lisa Kaati, and Ahmed Rezine. "Detecting key players in terrorist networks." 2012 European Intelligence and Security Informatics Conference. IEEE, 2012.

[136] Choudhary, Pankaj, and Upasna Singh. "A survey on social network analysis for counter-terrorism." International Journal of Computer Applications 112.9 (2015): 24-29.

[137] Pelofske, Elijah, Lorie M. Liebrock, and Vincent Urias. "Cybersecurity Threat Hunting and Vulnerability Analysis Using a Neo4j Graph Database of Open Source Intelligence." arXiv preprint arXiv:2301.12013 (2023).

[138] Winiecki, Donald, et al. "Validating bad entity ranking in the Panama Papers via open-source intelligence." 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE, 2020.

[139] Alotaibi, Norah, and Delel Rhouma. "A review on community structures detection in time evolving social networks." Journal of King Saud University-Computer and Information Sciences 34.8 (2022): 5646-5662.

[140] Huang, Shin-Ying, Yen-Wen Huang, and Ching-Hao Mao. "A multi-channel cybersecurity news and threat intelligent engine-SecBuzzer." Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2019.