OSINT investigativa
Lo stato dell'arte
Uliano Conti

1. Che cos'è l' OSINT

L'OSINT è il metodo più semplice per raccogliere informazioni, poiché è una forma di raccolta di dati attraverso fonti aperte (internet, trasmissioni, documenti, ecc.)[1].
OSINT è una parola composta: Open Source e Intelligence. Si riferisce al processo generale in cui chiunque può raccogliere e analizzare notizie basate su informazioni open source.
La parola Intelligence si riferisce alle informazioni e allo spionaggio. In particolare, si riferisce alle informazioni raccolte, elaborate e ridotte per soddisfare esigenze o renderle comprensibili. Open sources data si riferisce a dati generali, non trattati. Gli esempi di dati e informazioni includono immagini, fotografie, dati di indagine, audio, metadati e set di dati, che possono essere ottenuti da fonti pubbliche. Mentre con Open sources information si indicano i dati generali che sono stati parzialmente filtrati in base ad alcuni criteri. Gli esempi includono libri, articoli e articoli scritti su alcuni argomenti.
Uno dei passaggi più importanti di OSINT è la ricerca e la raccolta di informazioni pubbliche. Da un punto di vista tecnico l'OSINT mostra una serie di vantaggi, ma deve fare i conti anche con alcune limitazioni, che vengono descritte in dettaglio qui di seguito[2].

1- *Raccolta di informazioni in tempo reale*: le informazioni raccolte da OSINT sono rapidamente ottenute attraverso fonti aperte e i dati vengono tracciati in tempo reale;
2- *Acquisizione sicura di molti dati*: i dati raccolti da OSINT proteggono dati che implicano la presenza di informazioni segrete;
3- *Chiarezza delle fonti*: in HUMINT, la credibilità di dati è discutibile, mentre i dati raccolti da OSINT garantiscono maggiore credibilità;
4- *Convenienza e facilità di accesso*: non tutti possono accedere facilmente ai dati perché i diritti di accesso ai dati sono impostati in modo tale che possano accedervi solo gli utenti autorizzati;
5- *Basso costo*: OSINT ha il vantaggio di ottenere dati a basso costo, rispetto al costo della formazione degli agenti in HUMINT.
6- *Elevata capacità di calcolo*: i progressi dei computer consentono di eseguire operazioni ad alta intensità di lavoro in termini di raccolta, elaborazione, analisi e archiviazione.
Grazie a questa caratteristica, si ha l'opportunità di applicare OSINT prendendo in considerazione grandi quantità di informazioni pubbliche e di combinare un numero elevato di serie di dati, relazioni e modelli provenienti da diversi tipi di fonti aperte, applicando al tempo stesso tecniche avanzate di elaborazione e analisi.
7- *Big Data e apprendimento automatico*: vi è una continua proliferazione di tecniche di analisi e di data mining, nonché di algoritmi di apprendimento automatico che sono in grado di automatizzare e rendere più intelligenti ed efficienti i processi investigativi e decisionali. Ciò consente di individuare correlazioni complesse, naturalmente imprevedibili per gli esseri umani. Questo punto sarà fondamentale nelle future attività OSINT poiché segnerà la differenza tra la ricerca guidata dall'essere umano e quella guidata dall'intelligenza artificiale. Incorporando queste tecniche, il processo di raccolta e di analisi migliorerà, ottenendo così indagini più accurate e centrate all'obiettivo.

1 G. Nacci, "OSINT investigativa. Tecnologie e Analisi delle informazion", Intelligence & storia, Top secret 83-92, 2008
2 J. Pastor-Galindo, P. Nespoli , F. Gómez Mármol , and G. Martínez Pérez "The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends, IEEE Access (Volume: 8), 2020.

Inoltre, le agenzie di controspionaggio di un governo possono sfruttare tale paradigma per migliorare ulteriormente la qualità delle informazioni gestite e, di conseguenza, la lotta contro le organizzazioni criminale e terroristiche.
8- *Tipi di dati complementari*: possibilità di alimentare l'OSINT con altri tipologie di informazioni. La struttura intrinseca del sistema è sufficientemente aperta per includere dati che non sono stati effettivamente ottenuti da fonti aperte. Questo significa che l'OSINT può essere ancora più efficace se siamo in grado di aggiungere informazioni esterne che completano le indagini. Ad esempio, le forze dell'ordine potrebbero approfittare della collaborazione dei cittadini per alimentare le ricerche OSINT, i servizi di intelligence potrebbero sfruttare le informazioni classificate sui criminali informatici o sugli incidenti per arricchire le indagini OSINT.
9- *Scopo flessibile e ampio campo di applicazione*: grazie alla natura dell'OSINT, le indagini possono essere estese a molti problemi e possono raccogliere informazioni in tutto il cyberspazio. Questo paradigma può essere utilizzato per scopi economici, psicologici, strategici, giornalistici, lavorativi o di sicurezza. In particolare, si possono evidenziare i vantaggi nell'ambito della criminalità e della sicurezza informatica, dove l'OSINT potrebbe monitorare attori sospetti o gruppi pericolosi, individuare contesti e processi di radicalizzazione, studiare le tendenze preoccupanti della società, supportare l'attribuzione di cyberattacchi e crimini, o migliorare l'analisi forense digitale.

Gli svantaggi sono principalmente:
1- *Complessità della gestione dei dati*: la quantità di dati è enorme e, di conseguenza, è impegnativo gestirli in modo efficiente ed efficace. È bene per l'OSINT considerare il maggior numero possibile di informazioni, ma anche di disporre di tecniche avanzate e di risorse significative per garantire una raccolta, un'elaborazione e un'analisi di alta qualità. La quantità di informazioni è troppo vasta: più informazioni l'utente ha, e più è difficile produrre dati affidabili utilizzando OSINT. Inoltre se le informazioni errate sono mescolate ai dati, ciò può ridurre la credibilità dei dati. Attualmente, poiché molti dati vengono ricercati in fonti aperte, ci vuole tempo e fatica per rilevare informazioni false e selezionare dati affidabili.
2- *Percezione organizzativa e pregiudizio delle agenzie di intelligence*: nella cultura organizzativa delle agenzie di intelligence, il valore dei dati raccolti da OSINT è sottovalutato e l'importanza dei dati non è considerata perché chiunque può accedere e utilizzare i dati.
3- *Problemi di sicurezza e vincoli tecnici*: le agenzie di intelligence ricorrono a reti informatiche interne a causa di problemi di sicurezza, che limitano l'uso di OSD adoperando Internet. Di conseguenza, gli analisti delle agenzie di intelligence mostrano un atteggiamento passivo verso l'impego dei dati OSINT. Gli esperti di sicurezza informatica stanno cercando di preparare metodi per utilizzare liberamente OSD, risolvendo al contempo i problemi di sicurezza.
4- OSINT può diventare una pietra angolare dei crimini informatici quando viene utilizzata in modo improprio poiché chiunque può accedere ai dati raccolti dall'OSINT. Vi è lo svantaggio che i dati raccolti dall'OSINT possono essere la base per commettere crimini informatici. Pertanto, sono necessarie ulteriori ricerche sui requisiti di sicurezza e sulla tecnologia in grado di ridurre al minimo i danni dei crimini informatici, per evitare che gli utenti adoperino i dati OSINT per scopi dannosi.
5- *Informazioni non strutturate*: le informazioni pubbliche disponibili su Internet sono intrinsecamente e massicciamente disorganizzate. Ciò significa che i dati raccolti dall'OSINT sono così eterogenei che risulta difficile classificarli, collegare ed esaminare tali dati al fine di estrarre relazioni e conoscenze rilevanti. In questo senso, l'OSINT richiede meccanismi come il data mining, l'elaborazione del linguaggio naturale (NLP) o l'analisi del testo per omogeneizzare le informazioni non strutturate al fine di poterle sfruttare.
6- *La disinformazione:* i social network e i mezzi di comunicazione sono costellati da opinioni soggettive e fake news. Per questo motivo, l'esistenza di informazioni imprecise deve essere tenuta in considerazione nell'implementazione dei meccanismi dell'OSINT. Le attività dell'OSINT devono

sempre avere a che fare con informazioni affidabili e seguire linee di esplorazione fidate per garantire risultati positivi e convincenti.
7- *Affidabilità delle fonti di dati*: l'attendibilità e l'autorevolezza delle informazioni sono la chiave per il successo delle indagini OSINT. Idealmente, i dati raccolti dovrebbero provenire da fonti autorevoli, revisionate e fidate (documenti ufficiali, rapporti scientifici, mezzi di comunicazione affidabili). Ma nella pratica, nell'OSINT coesistono anche fonti soggettive o non autorizzate, come i contenuti dei social network, le fonti soggettive o non autorevoli, il contenuto dei social network o dei media manipolati. Anche se questo tipo di fonti è più incline alla disinformazione, è in realtà il luogo in cui è possibile estrarre più conoscenza per indagare su persone, gruppi o aziende.
8- *Forti considerazioni etico-legali:* con lo sviluppo dell'OSINT emergono numerose preoccupazioni in materia di privacy. Ci si interroga se l'OSINT costituisca o meno un problema etico nell'ambito dell'etica della raccolta di informazioni. Da un lato, anche se pubblicamente accessibile, l'OSINT ha il potere di divulgare informazioni che non sono esplicitamente pubblicate sul web. I risultati scoperti dovrebbero rispettare la privacy degli utenti e non rivelare questioni intime e personali, tenendo conto delle normative vigenti in materia (come il GPDR). In questo senso, aspetti come l'orientamento sessuale, il credo religioso, l'inclinazione politica o i comportamenti compromettenti possono essere presi da Internet, e questo processo di divulgazione può essere problematico in molti Paesi. D'altro canto, l'ambito delle ricerche basate su OSINT dovrebbe essere, per definizione, limitato alle fonti di dati aperte. In nessun caso i controlli di accesso o i metodi di autenticazione possono essere bypassati per estrarre conoscenza.
Recenti sviluppi per quanto riguarda Internet e big data hanno causato l'aumento della quantità di dati e accelerato l'avanzamento di intelligenza open-source, come appunto l'OSINT. Per questa ragione è necessario un filtraggio dei dati efficiente e affidabile, per estrarre i dati che l'utente desidera da OSINT.
Le organizzazioni o gli utenti utilizzano strumenti di automazione per filtrare i dati, in base allo scopo. Pertanto, è importante controllare continuamente lo strumento di automazione, lo standard per il filtraggio dei dati e sono necessarie ricerche sulla verifica dei dati estratti.
Rimane quindi una sfida per coloro che raccolgono e filtrano i dati da OSINT. Così come garantire la trasparenza dei dati: l'affidabilità dei dati raccolti è stata una questione critica per gli utenti che utilizzano i dati di OSINT.
È importante tenere un registro delle fonti dei dati estratti da OSINT. Gli utenti devono poter usufruire della trasparenza dei dati, e la ricerca su questo rimane ancora una sfida.
Vi è una mancanza di convalida delle procedure di gestione della privacy: molte aziende, come Facebook e Google, raccolgono molti dati dagli utenti online per l'intelligence commerciale. Nei dati raccolti online sono inclusi non solo i dati generali prodotti da utenti ma anche informazioni sensibili, come nomi, compleanni, indirizzi e numeri di passaporto. Molte aziende hanno rivelato che raccolgono e gestiscono i dati in modo anonimo, tuttavia, non è noto se questo venga fatto correttamente. Pertanto, la ricerca su privacy e procedure di gestione nei dati OSINT rimane ancora una sfida.

1.1. Le tecniche di analisi OSINT

Le principali tecniche di analisi dell'OSINT sono:
- *Analisi lessicale.* I dati grezzi devono essere esaminati per estrarre entità e relazioni dal testo. È essenziale applicare processi di traduzione alla lingua utilizzata nell'indagine OSINT.
I software sono, ad esempio, NVIVO e TalTac2.
- *Analisi semantica.* Avere un bagaglio di parole non è utile se non si estrae il significato. Con questo scopo di comprendere i dati, oggi si utilizzano algoritmi di elaborazione del linguaggio.

Inoltre, le tecniche di sentiment analysis consentono di contestualizzare post o opinioni soggettive per classificare lo stato emotivo dell'autore (ad esempio, positivo, negativo o neutro).
Infine, le procedure di scoperta della verità affrontano il difficile compito di risolvere le controversie nei dati provenienti da più fonti che si trovano su posizioni opposte sullo stesso argomento.
- *Analisi geospaziale.* Dati raccolti dai social network, eventi, sensori o indirizzi IP sono utili per essere analizzati da una prospettiva basata sulla localizzazione. In questo, l'utilizzo di mappe o grafici facilita la rappresentazione e la comprensione dei dati, oltre all'estrazione di connessioni significative tra eventi o persone.
Grazie, ad esempio, al sofware STATA e al modulo *spmap* di M. Pisati per STATA.
- *Analisi dei social media.* Le caratteristiche dei moderni social media consentono ai ricercatori di effettuare un'analisi degli utenti. In questo scenario, l'analisi dei dati sociali permette di creare una rete di contatti, interazioni, luoghi, comportamenti e gusti intorno al soggetto.
I risultati delle tecniche sopra citate sono considerati come informazioni in uscita e sono classificati in tre gruppi principali:
1- *Le informazioni personali*, che si ottengono principalmente dal nome reale, dall'indirizzo e-mail, dal nome utente, dai social network e dalla ricerca, indirizzo e-mail e dai motori di ricerca.
2- *Le informazioni organizzative* sono costituite da aspetti di un team o di un'azienda composta da individui. Sono raccolti essenzialmente mediante tecniche di reti sociali, motori di ricerca, localizzazione, nome di dominio e indirizzo IP.
3- *Le informazioni di rete* riguardano i dati tecnici dei sistemi e delle topologie di comunicazione, che di solito si ottengono attraverso le tecniche di localizzazione, nome di dominio e indirizzo IP.
Logicamente, questi tre blocchi di informazioni possono essere ampliati con altri elementi. Inoltre, una singola indagine può avere diversi tipi di informazioni in uscita che si intrecciano l'uno con l'altro.

2. Quando una fonte si definisce aperta e quale differenza c'è con l'informazione?

Ancor prima di abbracciare questa visione di OSINT, è necessario porsi una domanda di ordine assai più generale: chi o cosa è lecito considerare alla stregua di fonte?[3] Quanti e quali tra gli oggetti e i soggetti che appartengono ad un determinato dominio informativo sono effettivamente in grado di essere delle *fonti*, ovvero di esprimere una funzione epistemologica valida? E come si fa a riconoscerli?
Il problema però è che la vicinanza tra i concetti di "fonte" e di "informazione" è notevole e un modello di questo tipo può non essere sufficientemente rappresentativo per tutte le applicazioni. Per molti versi, infatti, la fonte è già "informazione" in quanto, per il semplice fatto che esiste, informa su molte cose.
Come si affronta la problematica? Come si discrimina se una fonte è tale (e cioè una entità che informa circa un contesto più ampio, al di fuori di sé) o se una fonte è solo informazione pura, cioè un'entità che si limita ad erogare il proprio personale carico contenuto semantico a prescindere dal contesto che la circonda?
Un metodo utile consiste nell'analisi della struttura del cosiddetto "carico pagante"[4], ovvero del carico di contenuto semantico. Se il *payload* è costituito in prevalenza dai cosiddetti "dati grezzi" allora si è potenzialmente in presenza di un oggetto che possiamo definire "fonte-informazione". Se, al contrario, la "potenziale fonte" presenta un carico pagante eterogeneo, che comprende sia "dati primari" che dati operazionali, metadati e dati derivati allora si è potenzialmente in presenza di un

3 G. Nacci, "Open souces intelligence abstraction layer", Edizioni Epokè, Novi Ligure (AL), 2014.
4 G. Nacci "Appunti sull'architettura sistemica delle fonti in OSIT", Fonti Aperte Etc. n 5/2017, scaricabile on line https://www.giovanninacci.net/blog/appunti-sulla-architettura-sistemica-delle-fonti-in-osint/

oggetto che possiamo definire “fonte-fonte” (o *fonte propria*), ossia è quella fonte che informa, oltre che su di sé, anche sul contesto in cui si trova. Quindi in una “fonte-fonte” oltre al carico pagante è sempre presente la descrizione di un sistema di relazioni significative che rimandano a conoscenze disponibili al di fuori di sé, relative ad altre fonti-fonti o informazioni[5]. Questa visione architetturale, statica, ben descrive l’aspetto strutturale ma è carente nella descrizione degli aspetti più prettamente funzionali e sistemici della fonte. Infatti, la sola circostanza di possedere una struttura di dati come quella appena enunciata non basta a fare di un oggetto una “fonte”.
Questo per anticipare il fatto che affinché un oggetto possa essere definito “fonte”, oltre ad essere “disponibile”, deve necessariamente possedere una serie di funzionalità finalizzate alla concreta accessibilità e fruibilità del proprio contenuto semantico. Senza di queste funzionalità l’oggetto rimane un oggetto come ogni altro, che si annuncia per il solo tramite di quelle sue proprietà fisiche che sono osservabili dall’esterno tramite percezione (dimensioni, posizione, parti, colore, materiale costruttivo, ecc.).
Memorizzazione, narrazione e socializzazione sono dunque le tre capacità che un modello di fonte deve necessariamente prevedere.
In conclusione, una fonte è un’entità informativa dotata dei tre sistemi: mnestico, narrante e relazionale, e che attraverso questi è in grado di interfacciarsi dinamicamente con il contesto in cui è incardinato, attivando scambi e trasferimenti informativi con l’ambiente (per quanto concerne in modo specifico all’Intelligence delle Fonti Aperte, con “ambiente” si fa riferimento al cosiddetto “network delle fonti”).
Una fonte aperta non è necessariamente “pubblica” e nemmeno per forza di cose “pubblicamente disponibile”. Una fonte aperta può tranquillamente essere “privata”, la maggior parte delle fonti lo è, e può tranquillamente non essere *pubblicamente* disponibile, anche se rimane certamente disponibile *al pubblico*.
Questo fraintendimento è in parte dovuto anche al fatto che la dottrina OSINT ha sempre passivamente accettato l’analogia tra “fonte aperta” e “fonte giornalistica”, un errore le cui conseguenze - in termini di mancato sviluppo della disciplina - l’OSINT subisce ancora oggi.
Fortunatamente esistono metodi più oggettivi, efficaci ed affidabili per valutare la *openness* di una fonte. In primo luogo, è necessario ricordare che la fonte, così come anche l’informazione, della quale rappresenta una classe speciale, nasce originariamente, naturalmente aperta.
Da ciò si deduce che è aperta quella fonte (o informazione) che non è ancora stata sottoposta ad un regime di “classifica”, ovvero un sistema che agisce sulle proprietà di disponibilità e accessibilità della fonte/informazione. Tali fonti e informazioni si definiscono con il termine inglese “unaffected” ovvero, inalterate, non influenzate, non compromesse, naturali. Dunque, più che parlare di fonti “aperte” o “riservate” la proposta è quella di dividere fonti e informazioni in *affected* e *unaffected*, a seconda del fatto che siano state sottoposte a una classifica di riservatezza o meno.
Se quanto fin qui detto può sembrare superficiale in relazione alla dottrina *convenzionale* dell’intelligence o dell’OSINT, sarà sufficiente ricordare alcune problematiche per dimostrare invece la grande rilevanza di questa visione sistemica della fonte.
Una di queste problematiche è il cosiddetto “diritto all’oblio”, ovvero quel diritto formalmente o legalmente riconosciuto agli individui di richiedere, a chi le ha pubblicate, la cancellazione delle informazioni che lo riguardano.
Tale diritto deriva da una recente evoluzione concettuale che riguarda la percezione dei *nostri* dati personali: si è passati da un vecchio concetto di “possesso” dei dati, dove i dati “appartengono” all’individuo allo stesso modo di come gli appartiene un oggetto fisico, ad una *filosofia dell’identità* dove “i miei dati sono io, mi costituiscono”.
Capire quindi il funzionamento dei sistemi di classifica di riservatezza e il modo in cui le fonti rispondono all’apposizione di una classifica modificando i propri assetti sistemici è cosa fondamentale ai fini della cosiddetta “validazione della fonte”.

5 G. Nacci, “Open sources intelligence abstraction layer”, Edizioni Epokè, Novi Ligure (AL), 2014.

Invece, per ciò che concerne l'informazione: qual è il reale livello di consapevolezza dell'oggetto "informazione" per chi si occupa di intelligence, a livello governativo o di business? Una non adeguata percezione della complessità della natura di questo elemento può compromettere la qualità dell'attività di intelligence e, di conseguenza, il valore stesso della funzione istituzionale dell'intelligence quale supporto alla capacità decisionale?
Per rispondere facciamo ricorso alla Teoria dei Livelli di Astrazione che è stata sviluppata dal filosofo italiano Luciano Floridi e ampiamente trattata nel suo recente "The Philosophy of Information".
Quando osserviamo un oggetto di solito tendiamo a non esplicitare le risposte relative al "che cosa" stiamo osservando in ciò che osserviamo. L'idea dei "livelli di astrazione" nasce dalla necessita di rendere esplicito il complesso dei dettagli, ovvero delle proprietà, che scegliamo di ritenere significativi e discriminanti in ciò che osserviamo, in altre parole ciò che per un determinato osservatore rende un "*qualcosa"* esattamente "*quella cosa*" e non altre.
Il Livello di Astrazione dell'OSINT può dunque essere interpretato come una delle possibili interfacce esistenti tra un sistema "S" qualsiasi e il suo ambiente. In questo senso, come tutte le interfacce, l'OSINT non genera o "inventa" informazione inedita. Quello che l'OSINT può fare è produrre la migliore conoscenza possibile per il sistema che sta interfacciando, fornendo modelli e metodi di rappresentazione capaci di evidenziare i livelli di veridicità dei contenuti semantici fattuali presenti all'interno dell'ambiente.
Questa funzione di interfaccia del livello di astrazione dell'OSINT si esplicita in due ruoli significativi: il primo consiste nella mediazione continua tra il livello di astrazione dell'intelligence convenzionale (ovvero del "sistema di intelligence") e la realtà (o il mondo) rispondendo cioè ad un'esigenza di indagine ontologica (cosa esiste nel mondo: fonti, informazioni, notizie, documenti, ecc.). Il secondo ruolo, maggiormente legato ad una visione sistemica interna, consiste nel rappresentare un potente strumento di indagine epistemologica sull'intelligence.
In questo senso l'interfaccia sarà tra l'intelligence (in quanto disciplina, attività e funzione) e gli studi di intelligence (in quanto conoscenza scientifica a disposizione su quella disciplina, su quelle attività e su quelle funzioni).
In questo dibattito si inserisce anche Giovanni Nacci, secondo il quale l'OSINT non è una mera *tecnica* o una *tecnologia,* ma è una disciplina nel senso pieno del termine: ovvero con un sistema di metodi, sistemi e prassi dotato di una propria epistemologia.
La visione di OSINT che porta avanti lo studioso, è la proposta di un'auspicabilmente e condivisa *Teoria Generale per l'Intelligence delle Fonti Aperte,* fondata sull'osservazione delle dinamiche e delle modalità attraverso le quali discipline anche molto distanti tra loro (filosofia, linguistica, scienze cognitive, logica, informatica, storiografia) prestano all'OSINT "blocchi epistemologici" più o meno estesi e articolati, nonché delle prassi attraverso le quali tali "blocchi" vengono incorporati da quelli già esistenti, concorrendo alla costruzione del bagaglio teoretico dell'OSINT (e di conseguenza degli *intelligence studies* in senso lato).
Questi "apporti" epistemologici avvengono all'interno di un sistema complesso, interattivo, dinamico e ampiamente riconfigurabile, dove però debbono necessariamente esistere precise *regole di produzione.* Dovendo ricorrere ad una "immagine mentale" per descrivere tale fenomeno, l'autore sostiene che le modalità costruttive attraverso le quali tali blocchi si assemblano siano in qualche modo analoghe a quelle che guidano le costruzioni dei mattoncini LEGO.
L'idea di base è che l'assetto teorico dell'intelligence delle Fonti Aperte possa essere paragonato ad una "libera" costruzione Lego, dove i mattoncini sono costituiti da definizioni, teorie, prassi, metodi, sistemi e tecnologie offerte da una pluralità di discipline, che vengono assemblati sulla base di regole sintattiche ("auto-esplicitate" dalla singola disciplina) e semantiche, in modo che la struttura risultante sia il più possibile conforme alle finalità ed agli scopi che si attendono dall'OSINT.
L'autore continua la riflessione descrivendo quelle che sono le caratteristiche della fonte/informazione:

*Narratività della fonte in OSINT***.** Sempre riguardo a ciò che differenzia l'informazione dalla fonte, il livello di astrazione dello storico/storiografico ci offre un approccio da una diversa angolazione. Per lo storico una delle caratteristiche osservabili principali che definisce la fonte è la sua particolare *funzione narrante.* Ciò presuppone l'esistenza di una stretta relazione "narrante/narrato", tra fonte e informazione. Questo è ciò che accade quando si dice che "un'informazione è stata segretata". In realtà l'informazione resta sempre la stessa, nulla è aggiunto (o tolto) alle sue proprietà specifiche. A variare, sia in senso qualitativo che quantitativo, è semplicemente la profondità e la conformità della narrazione che la fonte fa dell'informazione.
Nell'Open Source Intelligence lo studio della "narratività" è fondamentale per comprendere i meccanismi e le prassi che soggiacciono alle operazioni di *classifica*, ovvero all'implementazione di regimi di riservatezza delle informazioni. La "classifica" è una operazione che concerne la limitazione della disponibilità di un'informazione nei confronti di uno o più soggetti che hanno, o potrebbero avere, interesse a fruire di quella informazione. Questo vuol dire che la narrazione può essere effettuata, a discrezione della fonte e più o meno volontariamente, omettendo uno o più aspetti delle proprietà.
Ciò che è interessante notare è che il sistema di classifica non è applicato all'oggetto narrato (ovvero, in questo caso, all'informazione che infatti rimane sempre uguale a se stessa) ma bensì alla specifica azione narrante indirizzata ad un determinato corrispondente.
Non si può non osservare, infatti, quella che è definibile come una delle principali patologie dell'intelligence delle fonti aperte (almeno per come è stata fino ad oggi considerata la disciplina): la quasi totale dipendenza da fonti di prassi giornalistica (tanto nella forma convenzionale che nelle più recenti versioni *on-line*). Tale fenomeno, lo sbilanciamento su una unica tipologia di fonte, è assolutamente deleterio per l'OSINT perché presuppone un fraintendimento totale delle origini e delle finalità della disciplina stessa.
L'OSINT (e l'Intelligence in generale) non può e non deve sbilanciarsi passivamente solo sugli "oggetti notizia". Al contrario l'intelligence delle fonti aperte dovrebbe concorrere a scardinare questa "passività" generalizzata nei confronti dei modi di essere (e di fare) informazione. Per questo motivo sono importanti lo studio e l'analisi concettuale di questa realtà e dei fenomeni e delle relazioni che in essa si osservano.

3. Quali sfide di ricerca richieste per futuro sviluppo OSINT?

In conclusione, la quantità di dati generati dall'attuale mondo interconnesso è incommensurabile e gran parte di tali dati è pubblicamente disponibile, il che significa che è accessibile da qualsiasi utente, in qualsiasi momento, da qualsiasi punto di Internet. Open Source Intelligence (OSINT) è un tipo di intelligenza che in realtà beneficia di quella natura aperta raccogliendo, elaborando e correlando i punti dell'insieme cyberspazio per generare conoscenza. In effetti, i recenti progressi tecnologici e dei Big data, gli open data rappresentano una potente fonte di analisi per ottenere informazioni rilevanti sui comportamenti sociali ma anche offrendo nuove linee di azione contro le minacce informatiche e il crimine informatico[6].
Per quanto riguarda l'uso di OSINT per estrarre le opinioni sociali e le emozioni, Santarcangelo et al.(2015)[7] hanno proposto un modello per determinare le opinioni degli utenti su una determinata parola chiave attraverso i social network, studiando in particolare gli aggettivi e le negazioni usati nei tweet.

6 Y. Woon Hwang I. Yeong Lee H. Kim, H. Lee and D. Kim, "Cuuent status and security Trend of OSINT", Wireless Communications & Mobile Computing . 2/18/2022, p1-14.
7 V. Santarcangelo, G. Oddo, M. Pilato, F. Valenti, and C. Fornaro, ``Social opinion mining: An approach for Italian language,'' in *Proc. 3rd Int. Conf. Future Internet Things Cloud*, Rome, Italy, Aug. 2015, pp. 693-697.

Purtroppo, è una semplice soluzione basata su una parola chiave progettata solo per la lingua italiana, senza tener conto delle questioni semantiche. D'altra parte, il metodo proposto da Kandias et al. potrebbe mettere in relazione l'uso che le persone fanno dei social network (in particolare Facebook) e il loro livello di stress. Tuttavia, gli esperimenti sono stati condotti solo con 405 utenti, mentre al giorno d'oggi esiste la possibilità di elaborare una quantità di dati molto maggiori[8].
Un altro studio interessante è stato condotto in Sud Africa dove gli autori hanno applicato l'elaborazione del linguaggio naturale (Natural Language Processing - NLP) ai messaggi di WhatsApp al fine di prevenire il verificarsi di violenze di massa[9].
Purtroppo, l'indagine è limitata ai messaggi di testo, escludendo quindi informazioni vitali che possono essere divulgate attraverso materiale multimediale.
Nel contesto della criminalità informatica e della criminalità organizzata, ci sono diversi lavori che esplorano l'applicazione di OSINT per le indagini criminali. Ad esempio, OSINT potrebbe aumentare la precisione e l'accuratezza dei procedimenti giudiziari e degli arresti dei colpevoli con framework come quello proposto da Quick e Choo[10].
Concretamente, gli autori applicano l'OSINT ai dati forensi digitali di una varietà di dispositivi per migliorare l'analisi dell'intelligence criminale. In questo ambito, un'altra opportunità che l'OSINT offre è l'individuazione di azioni illegali e la prevenzione di crimini come attacchi terroristici, omicidi o stupri. Infatti, i progetti europei POOLICE e CAPER sono stati concepiti per sviluppare modelli efficaci di scansione automatica dei dati aperti al fine di analizzare la società e individuare la criminalità organizzata emergente[11].
In contrasto con i progetti precedenti, le cui proposte non sono state utilizzate in casi reali, Delavallade et al.[12] descrivono un modello basato sui dati dei social network in grado di estrarre indicatori futuri di criminalità. Tale modello viene applicato al furto di rame e alla propaganda jihadista.
Dal punto di vista della sicurezza informatica e del cyberdefence, l'OSINT rappresenta uno strumento prezioso per migliorare i nostri meccanismi di protezione contro gli attacchi informatici.
Hernández et al. propongono l'uso di OSINT nel contesto colombiano per prevenire gli attacchi e per consentire un'azione strategica. Il sistema include non solo plugin per la raccolta di informazioni, ma anche modelli di apprendimento automatico per eseguire l'analisi del sentiment.
Inoltre, il progetto europeo DiSIEM[13] si pone come primo obiettivo l'integrazione di diverse fonti di dati OSINT negli attuali sistemi SIEM (Security Information and Event Management) per aiutare a reagire alle vulnerabilità scoperte di recente nell'infrastruttura o addirittura di prevedere le possibili minacce emergenti. Inoltre, Lee e Shon[14] hanno progettato un framework basato sull'OSINT per ispezionare le minacce alla cybersecurity delle reti di infrastrutture critiche. Ciononostante, tutti questi approcci non sono stati applicati a scenari reali per cui la loro efficacia rimane discutibile.
Inoltre, una revisione sistematica degli approcci, delle metodologie e degli strumenti proposti dal mondo accademico dei dati disponibili al pubblico è stata condotta su 107 articoli di ricerca tra il

8 M. Kandias, D. Gritzalis, V. Stavrou, and K. Nikoloulis, ``Stress level detection via OSN usage pattern and chronicity analysis: An OSINT threat intelligence module," *Comput. Security*, vol. 69, pp. 317, Aug. 2017.
9 B. Senekal and E. Kotzé, ``Open source intelligence (OSINT) for conict monitoring in contemporary South Africa: Challenges and opportunities in a big data context," *Afr. Secur. Rev.*, vol. 28, no. 1, pp. 1937, Jan. 2019.
10 D.Quick and K.-K.-R. Choo, "Digital forensic intelligence: Data subsets and open source intelligence (DFINT+OSINT): A timely and cohesive mix," *Future Gener. Comput. Syst.*, vol. 78, pp. 558-567, Jan. 2018.
11 R. P. Pastor and H. L. Larsen, ``Scanning of open data for detection of emerging organized crime threatsThe ePOOLICE project," in *Using Open Data to Detect Organized Crime Threats*. Cham, Switzerland: Springer, 2017, pp. 4771.
12 T. Delavallade, P. Bertrand, and V. Thouvenot, ``Extracting future crime indicators from social media," in *Using Open Data to Detect Organized Crime Threats*. Cham, Switzerland: Springer, 2017, pp. 167-198.
13 *Diversity Enhacements for Security Information and Event Management Project*. Accessed: Jan. 9, 2020. [Online]. Available: http://disiemproject.eu/
14 S. Lee and T. Shon, "Open source intelligence base cyber threat inspection framework for critical infrastructures," in *Proc. Future Technol. Conf. (FTC)*, San Francisco, CA, USA, Dec. 2016, pp. 1030-1033.

2013 e il 2017[15] per discutere dello stato dell'arte della valutazione della veridicità, che nell'ultimo decennio è diventata una grande preoccupazione a causa della diffusione di fake news e deepfakes. In questa direzione, gli autori delineano la relativa immaturità di questo campo, individuando diverse sfide che caratterizzeranno le future tendenze della ricerca

4. Progetto di Piattaforma di Intelligence con strumenti OSINT e tecnologie Open Source

L'architettura delle infrastrutture digitali, sempre di più si basa su Internet, un ambiente non molto sicuro. Senza grandi progressi nella sicurezza di questi sistemi o la variazione significativa del modo in cui essi sono costruiti o gestiti, si mette in dubbio la possibilità di come le organizzazioni possano proteggersi dalla crescente minaccia della criminalità informatica. L'Italia non è immune da questa problematica, infatti l'infrastruttura digitale presenta delle vulnerabilità che hanno permesso ai criminali di violare i sistemi e le informazioni.
Attualmente l'attenzione si sta spostando sulle minacce alle infrastrutture critiche. Nel "nuovo" approccio dell'Intelligence di Stato si promuove la cultura dell'Intelligence Economica.
Il progetto "*Piattaforma di Business Intelligence*"[16] ha l'obiettivo di creare un supporto tecnologico alle operazioni di consulenza rispondendo a determinate specifiche funzionali: ritrovare e salvare informazioni OSINT mediante l'utilizzo di «keyword» ed opportuni altri parametri di ricerca.
Per rispondere in modo tempestivo ed affidabile si è progettata la Piattaforma avendo cura di includere tutti gli scenari di maggior interesse di business intelligence quali: security, web reputation, economic e travel security intelligence. Affinché possa essere ritenuta di ausilio, è necessario che la Piattaforma sia un sistema integrato di risorse sia proprietarie che open source e che tutto il processo possa essere monitorato e supervisionato dall'intelligenza umana.
La problematica da superare è che per potersi adeguare velocemente alle repentine variazioni derivanti dall'ambiente OSINT è necessario confrontarsi costantemente con le imprevedibili variazioni delle sorgenti da cui attingono i rispettivi dati. Di fatto queste variazioni non sono controllabili.
Il collegamento con le fonti di riferimento garantisce la possibilità di accedere a circa 220.000 notizie nuove pubblicate ogni giorno in circa 70 lingue diverse su una selezione di circa 7.000 siti di quotidiani di stampa nazionali, locali ed internazionali. Altre fonti informative sono i social network e i motori di ricerca. L'integrazione di queste informazioni con Twitter e YouTube permette di raccogliere e predisporre una mole di informazioni molto rilevante, per poi effettuare la relativa fase di analisi. Gli sviluppi futuri sono numerosi e di varia natura. In prima istanza è necessario il completamento dei requisiti funzionali per coprire completamente gli scenari di utilizzo previsti della Piattaforma di Business Intelligence. In ogni caso per poter utilizzare la piattaforma fuori da un contesto "prototipale" non è possibile prescindere dal fatto di adeguare la sua architettura rendendo disponibile la quantità di risorse hardware necessaria per garantire il suo funzionamento in sicurezza e con un adeguato livello di servizio. Inoltre, potranno essere integrate sulla base di specifici requisisti altre tipologie di fonti dati su cui effettuare integrazioni ed ulteriori analisi. Infine, l'implementazione di un sistema di valutazioni per le fonti e per le informazioni è un elemento necessario per l'aumento della qualità ed attendibilità delle informazioni rilasciate dalla piattaforma stessa.

Conclusioni

15 M. G. Lozano, J. Brynielsson, U. Franke, M. Rosell, E. Tjornhammar, S. Varga, and V. Vlassov, "Veracity assessment of online data," *Decis. Support Syst.*, vol. 129, Feb. 2020, Art. no. 113-132.
16 M. A. Brignoli e L. Franchina "Progetto di Piattaforma di Intelligence con strumenti OSINT e tecnologie Open Source", In Proceedings of the First Italian Conference on Cybersecurity (ITASEC17), Venice, Italy, 2017.

In conclusione, fin qui è stata data una panoramica dello stato dell'arte dell'OSINT oggi. Sono emersi come alcuni dei lavori sono discutibili a causa principalmente della loro scarsa applicazione in scenari reali. Manca ancora un approccio serio per trasformare l'OSINT in una soluzione robusta e autogestita. Tuttavia, si suggerisce l'integrazione di OSINT con altre fonti di dati, passando però per una migliore affidabilità e sicurezza nella protezione dei dati. A seconda dei dati disponibili e dell'obiettivo finale, una corretta selezione dello strumento più appropriato può fare la differenza. Tuttavia una combinazione variegata di questi strumenti è in realtà la chiave per ottenere risultati plausibili.
È importante segnalare come nel contesto della Spagna, le forze dell'ordine spagnole e i servizi di intelligence impiegano al loro interno procedure OSINT. Nonostante sia un aspetto riservato del loro funzionamento, l'OSINT è un elemento cruciale nel contesto delle loro indagini. Vale la pena sottolineare che la Spagna sarebbe un grande territorio dove ricercare, sviluppare e applicare questa metodologia grazie alla sua maturità Open Data. In realtà la Spagna è uno dei paesi più trasparenti d'Europa, secondo il Portale europeo dei dati.
Come future direzioni di ricerca, vi sono ancora alcune sfide aperte legate alla raccolta, all'analisi e all'estrazione di conoscenza reale proveniente da Internet. Aspetti come la disinformazione, la privacy e la legalità saranno importanti nel futuro di OSINT. C'è ancora molta strada da fare, per questo motivo la comunità dovrebbe affrontare le sfide discusse includendo tecniche avanzate e migliorare le prestazioni attuali. L'obiettivo finale dell'OSINT è di essere in grado di garantire la ricerca desiderata per un certo scopo, in modo automatizzato e auto-guidato.
Pensare che l'intelligenza possa sostituirsi a ogni altra procedura demandata alla sicurezza è un errore culturale.
Fare più intelligence non vuol dire ridurre le altre attività di sicurezza, anzi il contrario, perché sarà necessario un maggiore raccordo e coordinazione di tutte le altre istituzioni che a vario titolo si occupano di sicurezza.
L'intelligence non può essere usata come strumentalizzazione politica per mettere d'accordo tutti.
In certi ambiti accademici sono nate attività di ricerca sui possibili punti di contatto tra l'intelligence e l'investigazione al fine di trovare una migliore forma di integrazione.
Quindi in tutti quei contesti particolarmente complessi o dove ci sia scarsità di informazioni, i metodi e i sistemi di intelligence possono fare la differenza tra una decisione presa con un'intuizione e una fondata sulle possibili conseguenze che essa comporta.
Grazie all'intelligence si ha la possibilità di attingere a una rappresentazione chiara e dinamica e dettagliata di ogni elemento informativo. Un fattore molto importante poiché dopo ogni fase investigativa la quantità di dati prodotti è enorme ed è difficile a qualsiasi mente umana catalogare tutte queste informazioni. Infatti uno degli strumenti tecnologici e metodologici che più caratterizza l'Open Source Intelligence è il Text Mining, ovvero l'analisi e la comprensione automatica dei testi, indispensabile nell'attività investigativa in quanto permette la categorizzazione dei concetti espressi, e pertanto l'indicizzazione automatica dei fatti, luoghi ed individui in esso citati. L'Italia è tra i migliori a livello mondiale per l'implementazione di Sistemi di Open Source Intelligence orientato alla sicurezza.
Il 15 febbraio 2006 il Ministero dell'Interno ha lanciato un programma innovativo creando la prima stazione di polizia on line: *commissariatodips.it* che ha ricevuto diversi premi. Tale progetto si basa sulle tecnologie del text mining e fornisce funzionalità di ricerca e classificazione dinamica di informazioni provenienti da più fonti e permette di scoprire aspetti meno palesi delle informazioni a disposizione poiché consente di identificare relazioni tra oggetti diversi che altrimenti sarebbero rimasti latenti.
Avere una maggiore consapevolezza delle conoscenze a disposizione, consente di utilizzare al meglio le informazioni per prendere decisioni. In questo modo il progetto del Ministero dell'Interno

mette al centro il cittadino che così diventa sia fruitori di servizi (denunce on line, assistenza, etc.) sia come fornitore di informazioni.
È proprio la trasformazione del cittadino da semplice fruitore di sicurezza a soggetto proattivo integrato nel sistema informativo di sicurezza la più grande innovazione.
Già da diversi anni gli esperti si interrogano sul concetto di Citizen Intelligence, di come e in che modo il cittadino potrà diventare erogatore di sicurezza per sé e per la sua famiglia. L'esempio di *commisariatodips.it* è il primo passo verso l'integrazione collaborativa di metodi e sistemi, dove l'Italia risulta essere la prima.

Bibliografia

Delavallade, P. Bertrand, and V. Thouvenot, (2017) "Extracting future crime indicators from social media," in *Using Open Data to Detect Organized Crime Threats*. Cham, Switzerland: Springer.

Diversity Enhacements for Security Information and Event Management Project. Accessed: Jan. 9, 2020. [Online]. Available: http://disiemproject.eu/

Kandias M., Gritzalis D., Stavrou V., and Nikoloulis K., (2017) "Stress level detection via OSN usage pattern and chronicity analysis: An OSINT threat intelligence module," *Comput. Security*, vol. 69, pp. 317.

Lee S. and Shon T., (2016), "Open source intelligence base cyber threat inspection framework for critical infrastructures," in *Proc. Future Technol. Conf. (FTC)*, San Francisco, CA, USA.

Lozano M. G**.** , Brynielsson J., Franke U., Rosell M., Tjornhammar E., Varga S., and Vlassov V., "Veracity assessment of online data," *Decis. Support Syst.*, vol. 129, Feb. 2020, Art. no. 113132.

Quick D. and Choo K.-K.-R., (2018), "Digital forensic intelligence: Data subsets and open source intelligence (DFINT+OSINT): A timely and cohesive mix," *Future Gener. Comput. Syst.*, vol. 78.

Nacci G. "Appunti sull'architettura sistemica delle fonti in OSIT", Fonti Aperte Etc. n 5/2017, scaricabile on line https://www.giovanninacci.net/blog/appunti-sulla-architettura-sistemica-delle-fonti-in-osint

Nacci, G., (2014), "Open souces intelligence abstraction layer", Edizioni Epokè, Novi Ligure (AL).

Nacci G., (2008), "OSINT investigativa. Tecnologie e Analisi delle informazion", Intelligence & storia, Top secret 83-92.

Brignoli M. A. e Franchina L."Progetto di Piattaforma di Intelligence con strumenti OSINT e tecnologie Open Source", In Proceedings of the First Italian Conference on Cybersecurity (ITASEC17), Venice, Italy, 2017.

Pastor-Galindo J., Nespoli P., Gómez Mármol F., and Martínez Pérez G. (2020), "The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends, IEEE Access (Volume: 8).

Santarcangelo V., Oddo G., Pilato M., Valenti F., and Fornaro C., (2015)"Social opinion mining: An approach for Italian language," in *Proc. 3rd Int. Conf. Future Internet Things Cloud*, Rome, Italy, pp. 693-697.

Senekal B. and Kotzé E., (2019), "Open source intelligence (OSINT) for conict monitoring in contemporary South Africa: Challenges and opportunities in a big data context," *Afr. Secur. Rev.*, vol. 28, no. 1, pp. 1937.

Pastor R. P. and Larsen H. L., (2017), “Scanning of open data for detection of emerging organized crime threatsThe ePOOLICE project," in *Using Open Data to Detect Organized Crime Threats*. Cham, Switzerland: Springer, pp. 4771.

Woon Hwang Y., Yeong Lee I., Kim H., Lee H., and Kim D., (2022), “Cuuent status and security Trend of OSINT”, Wireless Communications & Mobile Computing . 2/18/2022, p1-14