Rivista di Criminologia, Vittimologia e Sicurezza

Organo ufficiale della
Società Italiana di Vittimologia (S.I.V.)

World Society of Victimology (WSV)
Affiliated Journal

Anno XV Gennaio-Dicembre 2021

Numero Unico curato dal prof. Luca Cimino, socio S.I.V.

Rivista di Criminologia, Vittimologia e Sicurezza

Rivista quadrimestrale fondata a Bologna nel 2007
ISSN: 1971-033X
Registrazione n. 7728 del 14/2/2007 presso il Tribunale di Bologna

Redazione e amministrazione: Società Italiana di Vittimologia (S.I.V.) - Via Sant'Isaia 8 - 40123 Bologna – Italia; Tel. e Fax. +39-051-585709; e-mail: augustoballoni@virgilio.it

Rivista peer reviewed (procedura double-bind) e indicizzata su:
Catalogo italiano dei periodici/**ACNP**, Progetto **CNR SOLAR** (Scientific Open-access Literature Archive and Repository), directory internazionale delle riviste open access **DOAJ** (Directory of Open Access Journals), **CrossRef, ScienceOpen, Google Scholar, EBSCO Discovery Service, Academic Journal Database, InfoBase Index**

Rivista di Criminologia, Vittimologia e Sicurezza

ISSN 1971-033X
Anno XV, Numero Unico
Gennaio-Dicembre 2021

INDICE

doi: 10.14664/rcvs/134

Intelligenza artificiale e machine learning: nuovi strumenti per il contrasto della conflittualita' asimmetrica e per la gestione delle crisi - il caso di studio pandemia covid-19

Intelligence artificielle et apprentissage automatique : de nouveaux outils pour contraster les conflits asymétriques et pour la gestion de crise - l'étude de cas sur la pandémie de covid-19

Artificial intelligence and machine learning: new tools for contrasting asymmetrical conflict and for crisis management - the covid-19 pandemic case study

Roberto Mugavero[*]*, William Thorossian*[**]

Riassunto
Conflitti, rischi asimmetrici, instabilità, criminalità e terrorismo sono fenomeni che, assieme a dirompenti eventi come la pandemia COVID-19 sempre più caratterizzano in modo globalizzato gli scenari della sicurezza internazionale. Questa nuova realtà richiede metodologie e soluzioni innovative volte a migliorare il rilevamento, la conoscenza e la comprensione dei fenomeni al fine di prevenire o rispondere a potenziali minacce a livello locale, regionale ed internazionale. Questo documento affronta il problema ed analizza come l'intelligenza artificiale può essere usata per migliorare le competenze di sicurezza attraverso lo sviluppo e l'uso di una piattaforma basata sull'intelligenza artificiale e sugli algoritmi di apprendimento automatico.

Résumé
Les conflits, les risques asymétriques, l'instabilité, la criminalité et le terrorisme sont des phénomènes qui, associés à des événements perturbateurs tels que la pandémie de COVID-19, caractérisent de plus en plus les scénarios de sécurité internationale de manière mondialisée. Cette nouvelle réalité nécessite des méthodologies et des solutions innovantes visant à améliorer la détection, la connaissance et la compréhension des phénomènes afin de prévenir ou de répondre aux menaces potentielles au niveau local, régional et international. Cet article aborde le problème et analyse comment l'intelligence artificielle peut être utilisée pour améliorer les compétences en matière de sécurité grâce au développement et à l'utilisation d'une plate-forme basée sur l'intelligence artificielle et les algorithmes d'apprentissage automatique.

Abstract
Conflicts, asymmetric risks, instability, crime and terrorism are phenomena that, along with disruptive events such as the COVID-19 pandemic, increasingly characterize international security scenarios in a globalized way. This new paradigm requires innovative methodologies and solutions aimed at improving detection, knowledge and understanding of phenomena in order to prevent or respond to potential threats at local, regional and international level. This paper addresses the problem and analyzes how artificial intelligence can be used to improve security skills through the development and use of a platform based on artificial intelligence and machine learning algorithms.

Key words: interoperabilità semantica, minacce asimmetriche, intelligenza artificiale, *machine learning*, criminalità, terrorismo, COVID-19.

[*] Università di Roma "Tor Vergata", Dipartimento di Ingegneria Elettronica – DIE, Università della Repubblica di San Marino, Centro per gli studi sulla Sicurezza – CUFS.
[**] Osservatorio sulla Sicurezza e Difesa CBRNe - OSDIFE

1.Introduzione

L'evoluzione degli scenari di rischio legati alle trasformazioni globali insieme a eventi naturali, incidenti, conflitti, instabilità, terrorismo e minacce, hanno favorito lo sviluppo di nuove conoscenze e tecnologie volte a migliorare la salute e il benessere dei popoli. Questi cambiamenti hanno richiesto necessariamente nuovi approcci e l'emergere di nuovi paradigmi. Un esempio importante è rappresentato dagli atti intenzionali o dalle minacce che comportano il rilascio deliberato di sostanze pericolose per causare danni. Queste sostanze pericolose possono includere sostanze chimiche, agenti biologici e materiali radiologici, ed esistono una varietà di mezzi e meccanismi, anche auto costruiti, che possono permetterne il rilascio nell'ambiente circostante in una varietà di forme.

Gli sviluppi scientifici e tecnici attuali e futuri avranno un impatto sulla protezione dei cittadini, dell'ambiente e degli interessi strategici dei paesi e delle comunità internazionali. In questo contesto, l'intelligenza artificiale può essere uno dei principali "game changer" per supportare gli analisti impegnati ad approfondire la conoscenza delle potenziali minacce e la comprensione del futuro.

Muovendo dalle attività di ricerca interne esistenti, l'Osservatorio sulla Sicurezza e Difesa CBRNe OSDIFE - Italia, in collaborazione con l'Università di Roma "Tor Vergata" - Dipartimento di Ingegneria Elettronica - Italia, l'Università Statale della Repubblica di San Marino - Centro Studi sulla Sicurezza, la Flinders University - Australia e stakeholder italiani, ha lavorato sulla verticalizzazione delle tecnologie di cognitive computing, basate su ontologie e algoritmi di Machine Learning (afferenti alla tecnologia Cogito®), personalizzando le soluzioni tecnologiche al dominio dei rischi asimmetrici, e con un particolare focus sperimentale sul fenomeno COVID-19.

Il paper presenta l'evoluzione di un'attività di ricerca e analisi esistente che è stata integrata attraverso la sperimentazione di una tecnologia AI based, proponendo infine un proof of concept di un lite tool che, con sviluppi personalizzati, può essere adottato a diversi livelli, anche dove sono disponibili meno competenze e risorse economiche (unità locali di organizzazioni complesse, pubbliche amministrazioni di paesi in via di sviluppo, PMI, ONG, media).

2.Dettagli sperimentali

2.1. Materiali e procedure

La capacità e l'abilità di contrastare il rischio di minacce asimmetriche, criminalità e terrorismo e di fornire contromisure possono essere aiutate dalla raccolta e dall'analisi dei dati.

Esempi di strumenti di raccolta dati sono il Global Terrorism Database (GTD), dell'Università del Maryland e l'Incident and Trafficking Database (ITDB), sviluppato dall'Agenzia internazionale per l'energia atomica.

Entrambi includono dati e informazioni sistematiche, così come aggiornamenti periodici su eventi terroristici e/o eventi legati a minacce specifiche a livello nazionale e internazionale. Tuttavia, nessuno dei due database si concentra principalmente o collettivamente su minacce asimmetriche, criminalità e terrorismo, o informa il contesto in cui tali rischi e il loro impatto possono essere analizzati.

Questa ricerca è stata innescata dalla necessità di migliorare l'esistente "Report on CBRNe Events in the World" mensile e il database lanciato dall'Osservatorio sulla sicurezza e la difesa CBRNe (OSDIFE) nel 2014.

L'Osservatorio sulla Sicurezza e Difesa CBRNe (OSDIFE), in collaborazione con i suoi partner italiani e internazionali, attualmente ospita un database che raccoglie i dati open source relativi agli attuali incidenti e minacce asimmetriche e la loro distribuzione geografica, fornendo newsletter e report sintetici agli utenti finali tra cui agenzie di sicurezza, università e organizzazioni internazionali.

Il team di ricerca dell'OSDIFE raccolgono manualmente i dati open source (attraverso l'uso di stringhe di ricerca sul web, e il monitoraggio di fonti qualificate) e inseriscono i dati nel database. Il rapporto è prodotto manualmente su base mensile da un team di analisti, prima di essere distribuito agli abbonati.

Ulteriori considerazioni sono state fatte in relazione all'uso delle piattaforme di social media (e il suo potenziale di doppio uso nello spazio delle minacce asimmetriche, criminalità e terrorismo), in quanto il team ha riconosciuto i limiti e la funzionalità dei metodi attuali utilizzati per curare il database e produrre i rapporti mensili. I post dei social media, che possono essere sfumati e contengono possibili indicatori delle minacce asimmetriche, può significare che il vero messaggio diventa difficile da rilevare, e che la sorveglianza di più di 1 miliardo di post di social media al giorno da solo su tre piattaforme di social media, non sarebbe possibile o praticabile. Per questo motivo, l'attenzione specifica è stata focalizzata sulla creazione di un sistema in grado di raccogliere contenuti da fonti aperte come Web, Blog, RSS e Social Network (Twitter, Facebook).

2.2. Materiali e metodi

In questo contesto, la sfida critica è quella di automatizzare la reportistica per supportare l'analisi di enormi quantità di dati in modo più efficace ed efficiente rispetto alle tradizionali tecnologie basate su parole chiave o statistiche.

L'opzione adottata all'inizio era l'uso di una serie di tassonomie combinate con un software che opera attraverso le varie piattaforme di social media per cercare parole o termini sfumati. Una tassonomia è definita come la scienza o la tecnica di classificazione, o una classificazione in categorie ordinate. Avere i mezzi tecnologici per creare termini di ricerca basati su una serie di tassonomie, permette ai ricercatori di personalizzare una ricerca su diverse piattaforme di social media open source per soddisfare l'obiettivo organizzativo. La capacità di creare rapporti in tempo reale efficaci ed efficienti, fornirà alle organizzazioni un vantaggio temporale vitale per neutralizzare o disinnescare potenziali minacce.

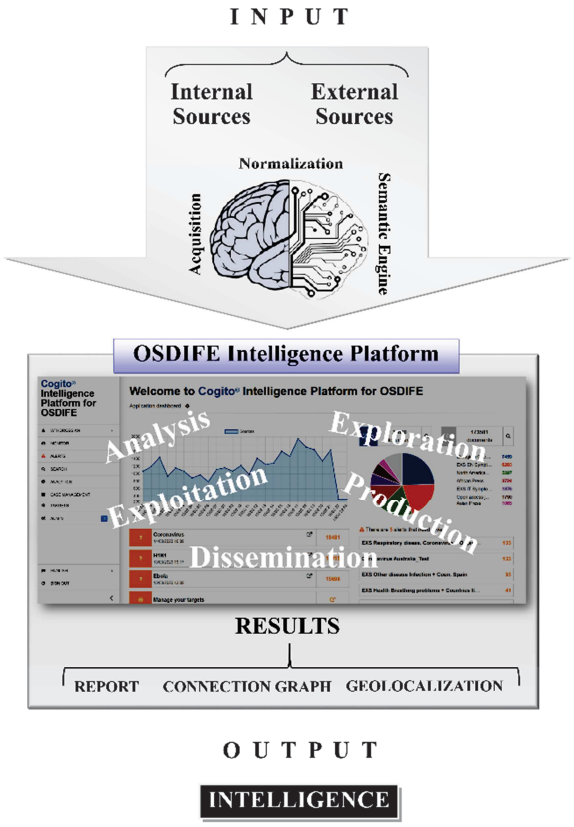

Fig.1. Schema funzionale della piattaforma di Intelligence OSDIFE basata sulle tecnologie Cogito®.

Nel caso specifico, il team OSDIFE ha cercato di aumentare la capacità del database di raccogliere, collegare e analizzare i dati relativi alle potenziali minacce asimmetriche, criminalità e terrorismo di interesse, per generare specificamente rapporti che possono fornire analisi delle tendenze, delle minacce e delle fonti di intelligence, con applicazione in tutto il mondo accademico, sanitario e della sicurezza. Il team ha intrapreso uno studio, progettato per valutare l'esperienza dell'utente finale del database e valutare le esigenze degli utenti come mezzo per informare un aggiornamento dello stesso database.

Lo scopo dello studio era di informare e sviluppare un crawler per contenuti strutturati e non strutturati basato sul software Cogito®, che, combinato con varie tassonomie, avrebbe fornito i requisiti di un sistema automatizzato per soddisfare le esigenze delle singole organizzazioni.

2.3. Fase preliminare

Il team ha coinvolto partecipanti e utenti finali, basati su organizzazioni che storicamente ricevono i rapporti OSDIFE come: agenzie di sicurezza internazionali e nazionali, organizzazioni non governative e accademiche; che lavorano nelle discipline della salute e della sicurezza. Lo studio ha utilizzato indagini per valutare la funzionalità e informare sugli aggiornamenti del database, scoprendo quali miglioramenti specifici al software potrebbero beneficiare per gli utenti finali nel tracciare, anticipare e prevedere meglio eventi e tendenze che altrimenti potrebbero non essere evidenti. Per valutare l'esperienza dell'utente finale e discernere le aree in cui il database potrebbe essere migliorato, è stato progettato un sondaggio di 20 domande per valutare come l'utente finale ha interagito con il database al momento, come vorrebbe vedere il database migliorato per fornire più informazione per l'intelligence o creare ulteriori collegamenti, e in quali modi un database migliorato aumenterebbe la sua analisi delle minacce asimmetriche, criminalità e terrorismo.

Il sondaggio è stato completato in due fasi. La prima fase è stata completata durante un incontro tenuto da OSDIFE a Roma. La seconda fase prevedeva un'indagine completata online utilizzando il software di indagine Qualtrics. I risultati di entrambe le indagini sono stati poi aggregati dal team. Una volta che i dati del pre-sondaggio iniziale sono stati analizzati, sono state redatte delle raccomandazioni per aumentare la funzionalità del database. Il team ha quindi lavorato per aggiornare il software e la capacità del database secondo le raccomandazioni fornite. Un sondaggio successivo è stato condotto dopo che i miglioramenti del database sono stati sviluppati e implementati per valutare il miglioramento delle funzionalità. Il post-sondaggio ha misurato la soddisfazione dell'utente finale per quanto riguarda gli aggiornamenti del database, e ha misurato se il miglioramento dell'usabilità è stato raggiunto in confronto con le risposte iniziali fornite nel pre-sondaggio.

3.Risultati e sviluppo

3.1. Soluzione tecnologica adottata

L'obiettivo principale è stato quello di fornire uno strumento in grado di raccogliere informazioni open source in un contesto di minacce asimmetriche/ criminalità/terrorismo, rafforzare in particolare i dati intorno agli incidenti biologici, e generare risultati che possono aiutare l'analisi delle tendenze, delle minacce e delle fonti di intelligence, con applicazione attraverso la sicurezza, il mondo accademico e i campi sanitari.

La soluzione IT sviluppata, offre supporto alla gestione della conoscenza del rischio delle minacce asimmetriche, criminalità e terrorismo, monitorando

una vasta gamma di fonti di informazione. A questo proposito, sono state create terminologie normalizzate, basate su un'ontologia opportunamente sintonizzata, e in grado di migliorare l'interazione e la comunicazione tra diverse entità internazionali (interoperabilità semantica).

Una piattaforma di Visual analytics, permette di:

- analizzare sistematicamente e continuativamente fonti di informazione online e offline come: Surface/Deep Web, Social Networks e Data Base;
- utilizzare una varietà di funzioni (glossario, tassonomia, georeferenziazione, filtri e correlazione) per valutare eventi, scenari, minacce e la loro evoluzione nello spazio e nel tempo;
- usare gli attributi semantici per scoprire contenuti, elementi, obiettivi e argomenti di interesse per categorie, entità, relazioni o concetti cluster (non solo parole chiave);
- navigare nelle informazioni geografiche estratte e le relazioni tra entità;
- inviare automaticamente "avvertimenti precoci".

In particolare, l'uso di algoritmi e regole avanzate di intelligenza artificiale (sia deep semantic che machine learning, vedere Fig.2) permette di:

a. estrarre categorie rilevanti provenienti da diverse tassonomie (Minacce asimmetriche, Salute, Terrorismo Intelligence, Crimine, Cyber);
b. estrarre tipi di entità "standard" (persone, organizzazioni, luoghi, date);
c. estrarre tipi di entità "dominio" legati agli ambienti Minacce asimmetriche, Salute, Intelligence, Terrorismo, Crimine e Cyber;
d. estrarre dati e informazioni relativi alle categorie Minacce asimmetriche, Salute, Intelligence, Terrorismo, Crimine e collegarli a persone, organizzazioni, luoghi e date;
e. estrarre relazioni tra entità standard e di dominio;
f. fornire il rilevamento del sentimento e delle emozioni;
g. creare cluster di contenuti uguali o simili.

Algoritmi e Text Mining lavorano insieme sul processo di analisi di testi scritti in linguaggio naturale, al fine di estrarre informazioni di alta qualità dal testo.

Ciò implica la ricerca di modelli interessanti nel testo o l'estrazione di dati da inserire in un database No-Sql. I compiti di text mining includono la categorizzazione del testo, il clustering del testo, l'estrazione di concetti/entità, la produzione di tassonomie granulari, l'analisi del sentimento, il riassunto dei documenti e la modellazione delle relazioni tra entità (cioè l'apprendimento delle relazioni tra entità nominate).

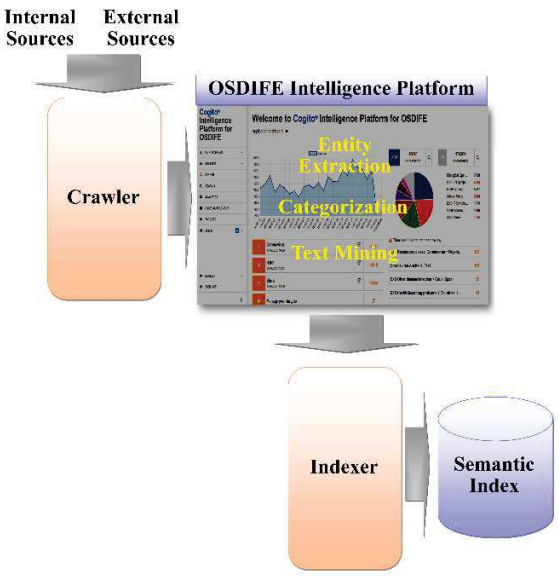

Fig. 2. Processo della piattaforma per catturare informazioni da varie fonti e creare un indice semantico

Gli sviluppatori devono preparare il testo usando l'analisi lessicale, il tagging POS (Parts-of-speech), lo stemming e altre tecniche di Natural Language Processing per ottenere informazioni utili dal testo.
Tecnicamente, l'estrazione e la classificazione efficace dei dati non strutturati, richiede un'analisi del testo e regole di gestione della classificazione su misura per le esigenze dell'organizzazione coinvolta e del settore in cui opera, così come i requisiti specifici del progetto (Fig. 3).

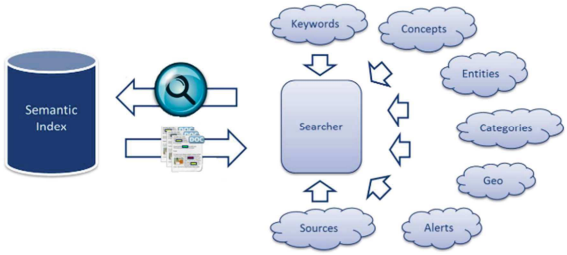

Fig. 3. Estrazione, indicizzazione e ricerca utilizzando l'indice semantico

I seguenti elementi sono stati adottati per sviluppare la piattaforma:

- Sviluppare regole di classificazione personalizzate.
- Utilizzare algoritmi di apprendimento automatico per estrarre le regole e automatizzare lo sviluppo.
- Un ambiente integrato per modellare l'analisi del testo e i progetti di arricchimento della conoscenza.
- Supporto multilingue e scalabilità.
- Gestione di progetti con diversi livelli di complessità.
- Importazione di tesauro e vocabolari.

Gli sviluppatori devono mettere insieme il contenuto testuale usando l'analisi lessicale, il tagging POS (parti del discorso), lo stemming e diverse strategie di Natural Language Processing per ottenere fatti utili dal contenuto testuale.

3.2. Casi d'uso

L'individuazione delle informazioni utili, avviene attraverso la corretta interrogazione della piattaforma, attraverso delle apposite "query", che permettono di filtrare, grazie alle capacità semantiche e la tassonomia correttamente implementata, la mole di dati che pervengono dalla ingestione dei vari contenuti dalle varie sorgenti, immesse attraverso un oculato processo di "source managment".

Di seguito vengono riportati alcuni casi d'uso dell'utilizzo della piattaforma, durante il processo di "tuning" e sviluppo, che hanno permesso di individuare, attraverso lo "scraping" da varie sorgenti, notizie e informazioni open source, atte all'individuazione precoce di possibili minacce nei teatri di maggiore interesse: Cyber crime, terrorismo, attacchi e guerriglie, correlazioni di malattie, COVID-19.

Individuazione di tentativi di attacchi informatici specifici su determinati protocolli ed in determinate aree geografiche (Fig. 4):

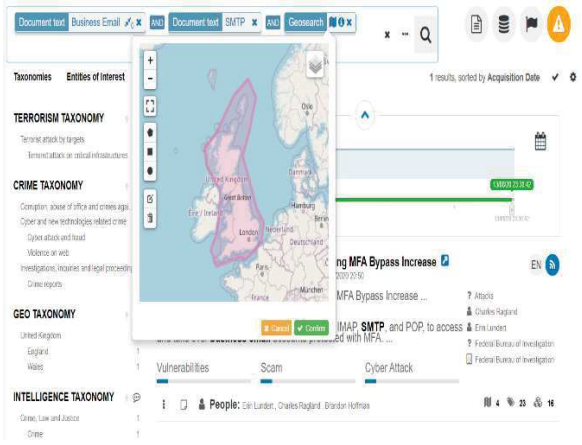

Fig. 4. Geo localizzazione di tentativi di Cyber crime su email con protocollo SMTP in Inghilterra

Analisi del "sentiment", che porta all'individuazione di possibili cyber attacchi in cui, risulti coinvolta una determinata entità (Fig. 5):

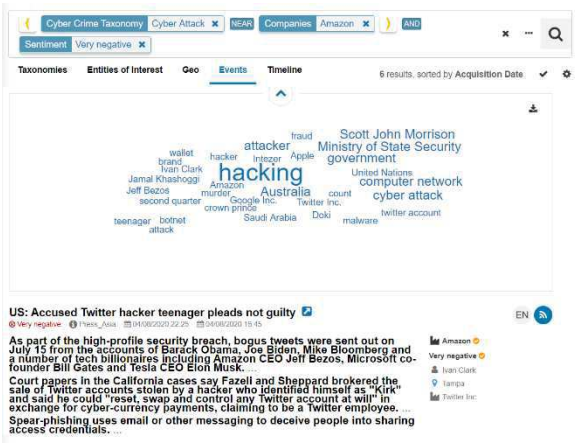

Fig. 5. Individuazione targettizzata su obiettivo, di attacchi informatici

Individuazione di possibili tentativi di reclutamento da parte di organizzazioni terroristiche in un determinato periodo temporale (Fig. 6):

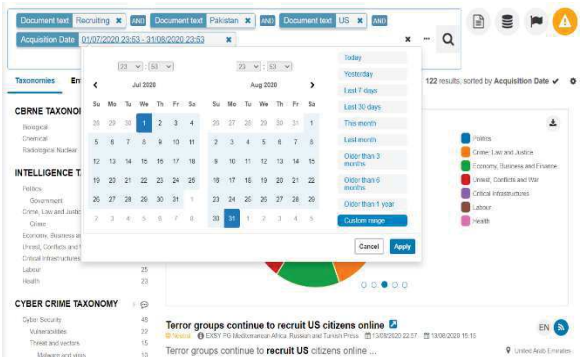

Fig. 6. Individuazione di possibili tentativi di reclutamento da parte di flange terroristiche Pakistane, di cittadini USA nel periodo 01/07/2020 fino al 31/08/2020

Reportistica di attacchi terroristici in specifiche aree e periodi (Fig. 7):

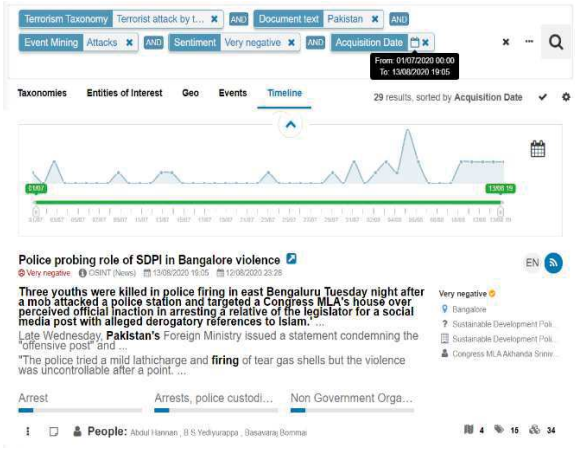

Fig. 7. Attacchi terroristici in Pakistan dal 01 luglio al 13 agosto 2020

Ricerca di articoli, sia scientifici che giornalistici, che trattino di terapie e cure da virus generici, e in cui sono coinvolti ricercatori in determinati campi ed in determinate aree geografiche (Fig. 8):

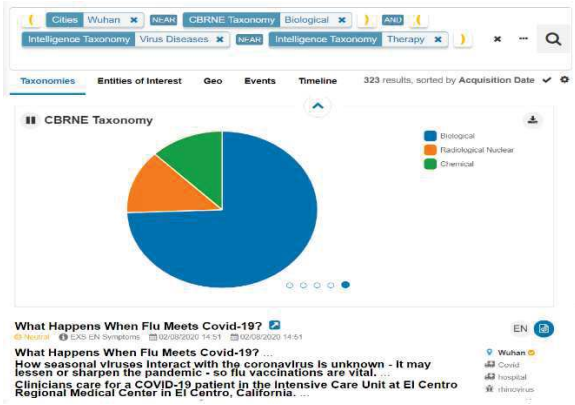

Fig. 8. Articoli che trattino di terapie e cure da virus che coinvolgano ricercatori in campo biologico della città di Wuhan

Analisi di informazioni relative alla produzione di un contromisure mediche in caso di pandemia (Fig. 9):

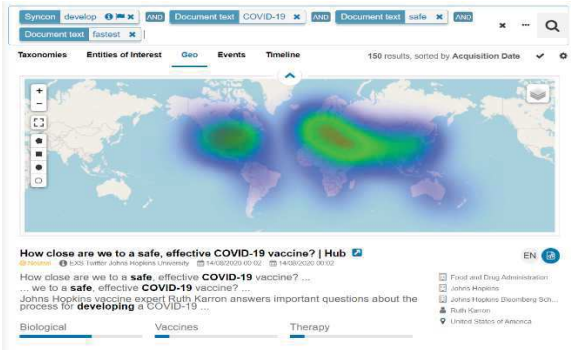

Fig. 9. Notizie che trattino degli eventuali inconvenienti nella produzione di un vaccino per COVID-19 in tempi veloci

Analisi del "sentiment" sulla gestione dell'emergenza COVID-19 (Fig. 10):

Fig. 10. "Sentiment" sulla gestione dell'emergenza COVID-19 in Italia

Identificazione di nessi causali tra malattie e sintomi, attraverso la connessione grafica di entità di interesse (Fig. 11):

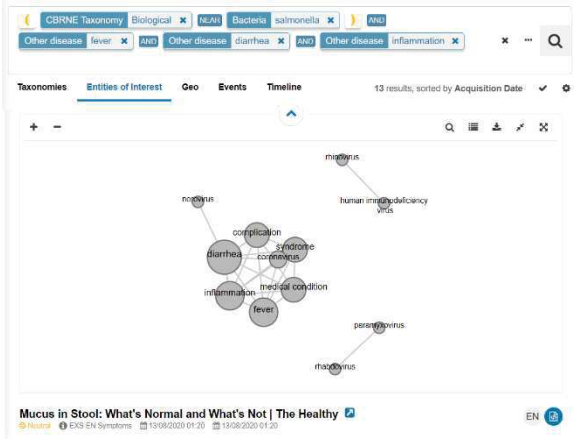

Fig. 11. Notizie, anche scientifiche, che riportino un nesso causale tra la salmonella ed alcuni sintomi

4.Conclusioni

Strumenti complessi e di alto livello sono necessari per la raccolta e l'analisi di informazioni "dall'alto verso il basso" ("top-down"). Solo lo "stato dell'arte" e le soluzioni di punta, possono essere adottate principalmente dalle amministrazioni centrali e dalle principali aziende.

Dall'altro lato, gli strumenti "lite" possono essere adottati a diversi livelli, anche dove sono disponibili meno competenze e risorse economiche (unità locali di organizzazioni complesse, amministrazioni pubbliche di paesi in via di sviluppo, PMI, ONG, media).

Per quanto riguarda quest'ultimo contesto operativo, i sistemi di monitoraggio e di allerta (ad esempio l'analisi del "sentiment" sui social media, il monitoraggio dei media locali e delle fonti dei social media) possono essere una svolta per gli attori locali, piccoli e medi che siano.

In tale contesto, la piattaforma ha manifestato una enorme utilità nel caso studio pandemico da COVID-19, rilevando, attraverso l'analisi dei social, la sua propensione all'individuazione di eventuali focolai in nascita. Tale analisi ci ha permesso di poter focalizzare un possibile utilizzo in tal senso, al fine di creare una sorta di "early warning" che possa essere utilizzato al fine di circoscrivere il rischio di contagi.

Strumenti flessibili possono essere rapidamente personalizzati per lavorare sia sul web che su dati statici (strutturati e non strutturati), e tale diffusione di strumenti più leggeri, può rafforzare la capacità di indagine del livello centrale, fornendo un'affidabile supporto dal basso verso l'alto ("bottom-up").

La soluzione informatica può permettere ad un team analitico di preparare rapporti relativi a specifiche domande di ricerca e analisi, legate alla valutazione delle minacce, all'identificazione delle

tendenze, al monitoraggio periodico di set di informazioni.

In questo quadro, l'uso effettivo della piattaforma è soggetto a un contatto preliminare tra l'utente finale e il team analitico, per definire la necessità. Infatti, per fornire un servizio di reporting dedicato, il team definirà, insieme all'utente, lo sforzo necessario per adattare le ontologie, per istruire la piattaforma con il know how semantico necessario per il dominio di studio richiesto, al fine di alimentare la piattaforma con le fonti necessarie.

Così, il futuro della piattaforma vede la realizzazione in due fasi:

- la ricerca di finanziamenti che permettano l'industrializzazione e la commercializzazione del prototipo pilota;
- la proposizione della piattaforma (Fig. 12) come un software fornito centralmente come servizio via Internet, cioè come un servizio Saas, "Software as a service".

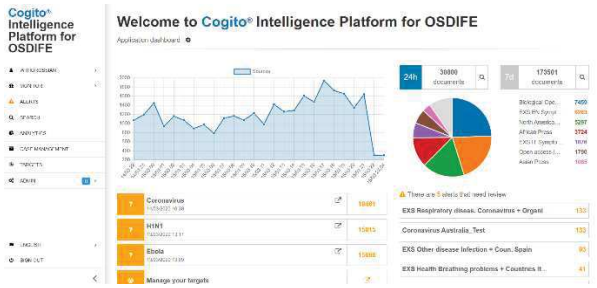

Fig. 12. Cruscotto della piattaforma di intelligence OSDIFE

Gli svantaggi e i possibili rischi di questo modello sono in gran parte di impatto limitato, anche considerando che il modello Saas si sta diffondendo rapidamente e la pressione competitiva sta contribuendo al continuo miglioramento della sicurezza dei dati e delle prestazioni.

L'uso delle sole ontologie nella navigazione dei database attraverso motori semantici è utile ma molto riduttivo. Le ontologie sono prima di tutto schemi e come tali hanno lo scopo di organizzare un dominio.

Questo può essere molto utile nei processi di produzione che coinvolgono una piccola comunità, dove un pezzo di informazione ha valore non solo in una fase del processo di produzione, ma può essere utile in diverse situazioni, ad esempio per un'ulteriore elaborazione. Più spesso, non è l'intera unità di informazione che è recuperabile, ma una parte di essa.

Questo richiede uno schema di organizzazione dei dati che possa dividere il dominio in tutte le classi di oggetti che giocano un ruolo nei processi. Le ontologie diventeranno probabilmente lo strumento più potente a disposizione di queste politiche di ricerca

semantica, e se si intende utilizzare Internet come infrastruttura, i linguaggi del Semantic Web diventano una risorsa applicativa indispensabile.

Anche se i livelli più alti dell'architettura del Semantic Web (Fig.13) possono richiedere diversi anni per raggiungere uno stadio in cui siano effettivamente implementabili e affidabili, c'è già una notevole quantità di lavoro nel settore delle ontologie.

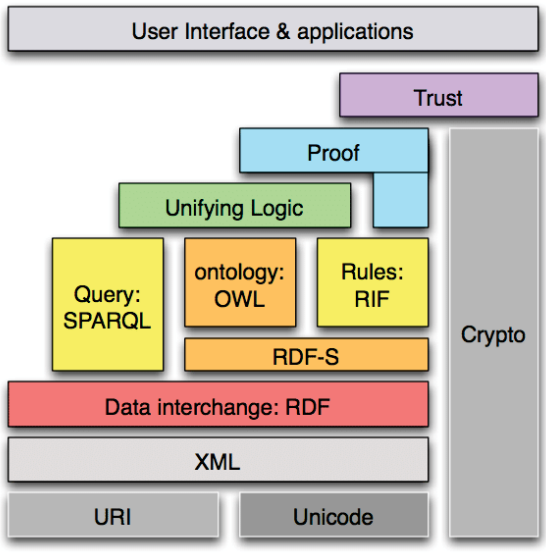

Fig. 13. Architettura del Web Semantico

Le soluzioni pratiche includono l'uso di XSLT (Extensible Stylesheet Language Transformations) per derivare RDF (Resource Description Framework) da documenti XML (Extensible Markup Language), l'emergere di database RDF (Resource Description Framework) generalizzati e motori di ricerca, interfacce grafiche generalizzate e specifiche RDF.

La prima cosa che dobbiamo quindi chiederci quando modelliamo una nuova ontologia è questa: cosa vogliamo dire, fare o chiamare i nostri oggetti? In alcuni casi potremmo aver bisogno di descrivere documenti che hanno a che fare con questi oggetti. In altri casi potremmo aver bisogno di tenere traccia di un processo di produzione. In altri casi potremmo aver bisogno di tenere traccia delle scelte e delle attività di un utente.

Il processo in cui lo schema deve essere inserito diventa il punto focale della modellazione. Se riusciamo a evidenziarlo chiaramente, sapremo quali oggetti descrivere. Di solito definiamo le ontologie di scopo come quelle che rappresentano la struttura dei processi. Le ontologie che forniscono gli oggetti specifici della nostra applicazione sono chiamate ontologie di dominio. Queste ultime sono quelle implementate nella piattaforma di intelligence OSDIFE, poiché lo scopo è quello di effettuare ricerche su specifici domini di interesse, per estrarre informazioni e notizie dalle fonti OSINT.

Bibliografia

- de la Torre-Abaitua G., Lago-Fernández L. F. and Arroyo D., A compression based framework for the detection of anomalies in heterogeneous data sources, *ArXiv,* vol. abs/1908.00417, 2019.
- Deliu I., Leichter C., and Franke K., Extracting cyber threat intelligence from hacker forums: Support vector machines versus convolutional neural networks, *IEEE International Conference on Big Data (Big Data)*, Dec 2017, pp. 3648–3656.
- Fayyad U., Piatetsky-Shapiro G. and Smyth P., From Data Mining to Knowledge Discovery in Databases, *AI Magazine* Volume 17 Number 3, 1996 (© AAAI).
- Ghazi Y, Anwar Z., Mumtaz R., Saleem S., and Tahir A., A Supervised machine learning based approach for automatically extracting high-level threat intelligence from unstructured sources, *International Conference on Frontiers of Information Technology (FIT)*, Dec 2018, pp. 129–134.
- Kim N., Lee S., Cho H., Kim B. and Jun M., Design of a cyber threat information collection system for cyber attack correlation, *International Conference on Platform Technology and Service (PlatCon),* Jan 2018, pp. 1–6.
- Lindstrom M., *Small Data: The Tiny Clues That Uncover Huge Trends*. Publisher: St. Martins Pr, Series: St. Martin's Press; ISBN-10: 1250080681
- Liu B. and Zhang L., *A Survey of Opinion Mining and Sentiment Analysis*. Boston, MA: Springer US, 2012, pp. 415–463.
- Liu B., *Sentiment Analysis*, Cambridge University Press, 2015, ISBN:9781139084789,vol. 203, pp. 91-98, 2008.
- Pellet H., Shiaeles S. and Stavrou S., Localising social network users and profiling their movement*, Computers & Security,* vol. 81, pp. 49 – 57, 2019.
- Ranade P., Mittal S., Joshi A., and Joshi K. Using deep neural networks to translate multi-lingual threat intelligenc, *IEEE International Conference on Intelligence and Security Informatics* (ISI), Nov 2018, pp. 238–243.
- Ruan D., Chen G.,. Kerre E. E, Wets G., *Intelligent Data Mining: Techniques and Applications (Studies in Computational Intelligence)*. Publishing house Springer Science & Business Media; ISBN 13: 9783642065767.
- Ruan. D, Chen G., Kerre E E., Wets G., *Intelligent Data Mining: Techniques and Applications (Studies in Computational Intelligence)* Springer Science & Business Media; ISBN 13: 9783642065767.
- Serrano L., Bouzid M., Charnois T., Brunessaux S. and Grilheres B., Events extraction and aggregation for open source intelligence: from text to knowledge, *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, pp. 518–523, 2013.
- Stieglitz S., Mirbabaie M., Ross B., and Neuberger C., Social media analytics - challenges in topic discovery, data collection, and data preparation, *International Journal of Information Management*, vol. 39, pp. 156 – 168, 2018.

- Vopham T., Hart J. E., Laden F., and Chiang Y. Y., Emerging trends in geospatial artificial intelligence (geoAI): Potential applications for environmental epidemiology, *Environmental Health*, vol. 17, no. 1, apr 2018.
- Wang M., Tsai M., Yang W. and Lei C., Infection categorization using deep autoencoder, *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, April 2018, pp. 1–2.
- Wang R., Ji W., Liu M.,. Wang X.,. Weng J, Deng S., Gao S. and an Yuan C., Review on mining data from multiple data sources, *Pattern Recognition Letters*, vol. 109, pp. 120 – 128, 2018, special Issue on Pattern Discovery from Multi-Source Data (PDMSD).

Sitografia.

- *COGITO® Intelligence Platform* (EXPERT.AI):https://www.expert.ai/de/resource/cogito-intelligence-platform-know-your-customer-better/
- *Global Terrorism Database (GTD), Maryland University*, https://www.start.umd.edu/gtd/
- *Incident and Trafficking Database (ITDB)*, IAEA International Atomic Energy Agency, https://www.iaea.org/resources/databases/itdb.