In Proceedings of the First Italian Conference on Cybersecurity (ITASEC17), Venice, Italy.

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

Mauro Alberto Brignoli [Vitrociset] e Luisa Franchina [Hermes Bay]
ma.brignoli@vitrociset.it, blustarcacina@gmail.com

Abstract

In questo documento sono presentati i temi di maggior rilievo ed interesse riguardanti gli aspetti di progettazione e realizzazione di una Piattaforma di Business Intelligence. Ne è riportata l'intera esperienza di progetto dalla definizione degli obiettivi di business con la scelta di utilizzare componenti open source per mantenere il costo totale il più contenuto possibile, alla formalizzazione dei requisiti ed infine all'implementazione avvalendosi di un team di sviluppo off-shore. Particolare attenzione è stata posta nell'organizzazione del team e alla comunicazione tra committente, analisti, project manager e sviluppatori implementando a tal proposito uno specifico meta-linguaggio per garantire l'integrità delle informazioni durante l'intero processo di analisi. Sono di seguito riportati il problema affrontato, i risultati raggiunti, le componenti innovative realizzate, le problematiche incontrate, lo stato dell'arte dei lavori nonché i possibili sviluppi futuri.

1 Il Problema

L'architettura delle infrastrutture digitali, sempre di più si basa su Internet (Confindustria, 2011), ambiente non 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 a questa problematica, infatti l'infrastruttura digitale presenta delle vulnerabilità che hanno permesso a criminali di violare i sistemi e le informazioni in essi contenute (CLUSIT, 2012 - 2013).

Attualmente l'attenzione si sta spostando sulle minacce alle infrastrutture critiche (COPASIR, 2010). Nel "nuovo" approccio dell'Intelligence di Stato si promuovere la cultura dell'Intelligence Economica. Si evidenziano necessità consulenziali in Business Intelligence (BI) con necessità di sistematizzare i processi di ricerca, l'analisi, la reportistica OSINT, finalizzando ad un risparmio di costi, a maggior tempestività nella risposta, un incremento dell'affidabilità complessiva del reporting e la riduzione delle polarizzazioni informative dovute alla mancata consultazione o valorizzazione di dati esistenti.

2 Progetto “Piattaforma di Business Intelligence”

Il progetto 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; arricchire le informazioni ritrovate mediante l’utilizzo di strumenti di analisi; trasformare, analizzare dati ed informazioni ritrovate, applicando algoritmi di classificazione, di clustering e regressione; visualizzare dati ed informazioni con report e dashboard dinamiche ed interattive; Capacità di memoria nelle operazioni svolte.

2.1 Requisiti di progetto

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**. Lo strumento deve essere progettato per agevolare e risolvere i problemi dal punto di vista dell’analista.

L’architettura deve utilizzare tendenzialmente tecnologia **open source** e deve avere tutte le componenti tecnologiche **virtualizzate** in modo da poter adottare il paradigma del cloud computing e quindi essere scalabile rispondendo alla necessità di un impiego di un numero di risorse adeguato e proporzionato alle esigenze operative (autores, 2013).

2.2 Comunicazione e modalità di interazione nel team di progetto

I componenti del team di progetto hanno prevalentemente interagito da remoto mediante l’utilizzo di strumenti di comunicazione online. In Figura 1 sono rappresentare le relazioni di prossimità fisica dei vari componenti del gruppo impegnato nella realizzazione del progetto. Il team leader fornisce i requisiti di business; il project manager coordina il gruppo di lavoro locale composto essenzialmente da analisti e quello di sviluppatori off-shore mantenendo la responsabilità e la gestione dello stato di avanzamento dei lavori di progetto. Per gestire la collaborazione remota la scelta è stata di un ambiente collaborativo di tipo cloud che implementa un protocollo di comunicazione “ad hoc”, brevemente descritto qui nel seguito. Per prima cosa si è provveduto a formalizzare i requisiti di progetto sotto forma di “**user story**”. Una “user story” è una descrizione della singola funzionalità che l’utente finale (analista) si aspetta di poter utilizzare una volta che questa sia stata realizzata. Le “user story” sono state raggruppate in gruppi di operazioni attinenti come la form di ricerca per esempio oppure la visualizzazione dei risultati ottenuti da una ricerca. A ciascuna “user story” è stato dato uno specifico **ID** ed alcuni attributi descrittivi per poterla meglio qualificare e gestire più agevolmente. Inoltre sono stati aggiunti degli attributi specifici utilizzati dal team durante l’implementazione della stessa. Il primo di questi attributi è **Validation** utilizzato dagli analisti e project manager per dare indicazioni agli sviluppatori. Lo stato “todo” si riferisce ad una “User Story” in attesa di essere realizzata; Una è volta che questa sia stata implementata correttamente lo stato passa ad “OK”; quando invece non fosse

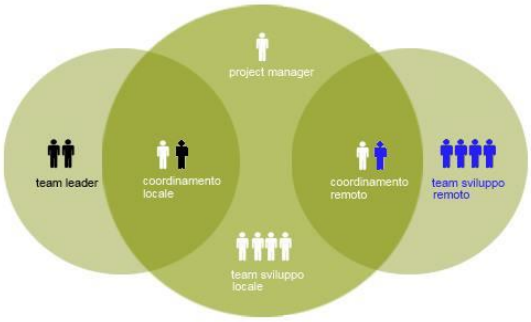

Figura 1 - Composizione del Team

ID	Name	Category	Desc	Analysts Comments	Validation (OK, todo, ?, KO)	Priority/ Severity (High, Medium, Low)	Developers Comments
1	cluster selection	mandatory	user select a cluster form a defined menu	cluster-dbmust be master detail	OK		
2	db selection	mandatory	users can select from db list menu for that particular cluster		OK		
3	name	mandatory	name of the query		OK		
4	source type	mandatory (1)	users can define a multiple sources selection: newspapers, social networks, blogs and forum, search engines (1) at least one value selected.	Search Engines and Blog & Forum are not available	KO	High	Fixed in new version

Figura 2 - User Story

corretta lo stato viene impostato a "KO"; lo stato "?" implica la necessità di rivedere la definizione da parte degli analisti per una sua migliore descrizione e/o formalizzazione. Il secondo attributo è **Priority/Severity**, anch'esso utilizzato da analisti e project manager per indirizzare le attività di sviluppo in una direzione piuttosto che in un'altra. Gli attributi **Analyst e Developers Comments** sono spazi per la comunicazione reciproca che creano un passaggio di informazioni utili e contestualizzate per permettere a ciascuno di portare avanti il proprio lavoro nelle rispettive prerogative. In Figura 2 è possibile vedere un estratto di una "User Story". L'utilizzo della piattaforma cloud collaborativa mette a disposizione degli strumenti come i commenti firmati e dotati di timestamp e la funzionalità di versioning del documento in modo tale da garantire sicurezza e flessibilità al protocollo adottato.

3 Risultati raggiunti, problematiche e possibili sviluppi

La composizione ed organizzazione del team è stata fondamentale per la progettazione architettonica della piattaforma e la gestione della complessità nell'integrazione dei dati, garantendo la scalabilità del sistema. A tal proposito è stato studiato ed adottato un **protocollo di comunicazione** tra analisti e sviluppatori specifico per il processo di definizione delle informazioni rilevanti che devono essere

ID	[Campo]	Included	Name	API attribute	API Desc
1	Identificativo	X	id	id_str	
2	Username [Met]	X	username		
3	Topic Principale [Sem]	Null			
4	Titolo	Null			
5	Sottotilo [Met]	Null			
6	Data	Null			
7	Content [Met]	X	con		
8	Luoghi(Paesi e città) [Met	X	plac		
9	Coordinate	X	Coo		
10	Entità (Compagnie, Persor	X	enti		

Figura 3 - Website Scraping Mapping

estratte da una pagina web di un generico sito (forum, blog etc.). Questo processo parte dalle indicazioni che l'analista fornisce allo sviluppatore, per esempio quali informazioni sono rilevanti all'interno di una pagina ed a quali attributi esse afferiscono. Ci si riferisce a questo processo con il termine anglosassone "web scraping" quando è automatica. Il processo si conclude con la verifica a posteriori che quanto implementato dallo sviluppatore corrisponda effettivamente a quanto voluto dall'analista. Nella Figura 3 è possibile identificare la tecnica utilizzata per questo passaggio di conoscenza.

L'**interfaccia** è di tipo web (Bookity, 2015) ed è un modulo completamente progettato e realizzato "ad hoc".

Il processo di analisi e gestione **dell'informazione** è stato scomposto nel dettaglio mostrando ciascuno dei passi significativi e necessari per il trattamento dell'informazione al fine di ottenere la Business Intelligence.

Nel diagramma di Figura 4 sono evidenziati gli elementi aggregati costituenti la "Base di Conoscenza" e le fonti informative complementari come le banche dati proprietarie.

La parte innovativa del progetto non è relativa ad una componente tecnologica quanto piuttosto alla parte di progettazione del processo di trattamento del dato partendo dalla sua fonte.

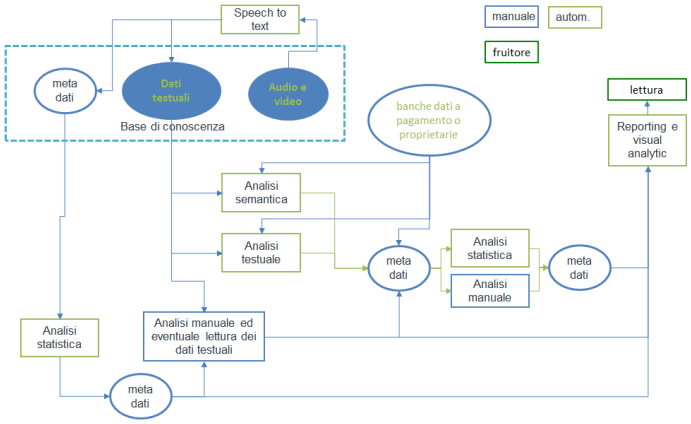

Figura 4 - Gestione informazione: Analisi

E' stata realizzata una "**mappatura**" di tutti gli attributi necessari per eseguire operazioni di Business Intelligence rispetto aii dati provenienti da ciascuna fonte considerata in modo tale che in fase di memorizzazione delle informazioni venisse conservata e garantita la coerenza semantica tra campi analoghi provenienti da sorgenti diverse. Nella Figura 5 è possibile visualizzare tale risultato identificando nelle colonne le sorgenti analizzate. In ogni riga sono rappresentati gli attributi ritenuti necessari a rispondere agli obiettivi di business posti. In ogni cella corrispondente all'incrocio tra righe e colonne è riportata una "x" quando è possibile riscontrare la presenza di tale attributo, mentre è presente il valore "Null" quando questo non è disponibile.

		Twitter	Facebook	YouTube	News	google	Search Engine	Blogs & Forum	sites
Column_id	DB_Column_name	IDT	IDF	IDY	IDN	Include	IDSE	IDBF	Included
1	id_result	X	X	X	X	X	X	X	X
2	username	X	X	X	Null	Null	Null	X	X
3	principal_topic	Null	Null	Null	X	X	X	X	X
4	title	Null	Null	X	X	X	X	Null	X
5	subtitle	Null	Null	Null	X	X	X	Null	X
6	data	Null	Null	Null	X	X	X	Null	Null
7	content	X	X	X	Null	Null	Null	X	X
8	place	X	X	X	X	X	X	X	X
9	coordinates	X	X	X	X	X	X	X	X
10	entity	X	X	X	X	X	X	X	X
11	relation	X	X	X	X	X	X	X	X
12	target	X	X	X	X	X	X	X	X
13	event_type	X	X	X	X	X	X	X	X
14	sentiment	X	X	X	X	X	X	X	X
15	link	X	X	X	X	X	X	X	X
16	pulishDate	X	X	X	X	X	X	X	X
17	source	X	X	X	X	X	X	X	X
18	source_trust	Null	Null	Null	X	X	X	Null	Null
19	comments_number	X	X	X	X	null	X	X	null
20	comments_positive	X	X	X	X	null	X	X	null
21	comments_neutral	X	X	X	X	null	X	X	null
22	comments_negative	X	X	X	X	null	X	X	null
23	repost_retweet	X	X	X	X	null	X	X	null
24	likes	X	X	X	X	null	X	X	null
25	popolarity_post-retweet	X	X	X	Null	Null	Null	X	null
26	views	Null	Null	Null	X	X	X	Null	Null
27	popolarity_news	Null	Null	Null	X	X	X	Null	Null
28	severe_index	X	X	X	Null	Null	Null	X	Null
29	language	X	X	X	X	X	X	X	X
30	likes_not	X	X	X	X	X	X	X	X
31	channel	X	X	X	X	X	X	X	X
32	query_id	X	X	X	X	X	X	X	X

Figura 5 - mappatura campi delle diverse fonti

3.1 Problematiche e aspetti critici di progetto

Nonostante il requisito di scalabilità funzionale fosse previsto e quindi implementato nei moduli di ricerca della piattaforma IMP, 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. Ne segnaliamo due tipologie diverse in cui ci siamo imbattuti.

Laddove è stato possibile si sono utilizzate le API native messe a disposizione della singola fonte senza utilizzare alcun "workaround" per recuperare i dati in forma testuale e associare tali informazioni nel database in corrispondenza dei rispettivi attributi, tuttavia, le API, quando disponibili sono in costante evoluzione e/o manutenzione evolutiva e non di rado nell'arco di poco tempo bisogna riadattare parte del codice scritto per potersi riadeguare alle modifiche fatte dal fornitore/titolare della fonte stessa.

Per la grande maggioranza dei casi l'acquisizione dei dati avviene mediante la singola lettura e memorizzazione delle pagine Internet. L'utilizzo del "**web scraping**" **è ancora più sensibile** alla variazione della struttura della pagina e questa variazione in termini di frequenza è maggiore rispetto alle variazioni riscontrate nelle API. Infatti, basti pensare che anche semplici aggiornamenti di tipo grafico, come un cambio di layout, possono determinare la "rottura" del codice utilizzato per estrarre le informazioni di interesse determinando l'interruzione della funzionalità di acquisizione dati relativamente alla fonte in oggetto. Inoltre il passaggio di conoscenza realizzato utilizzando la tecnica in Figura 3 è indipendente dalla tecnologia utilizzata ed è sempre soggetto a dei compromessi che possono preludere a vere e proprie difficoltà.

Il processo di **testing e verifica** è essenziale e sostanziale nel garantire un pieno controllo della direzione degli avanzamenti del progetto. Questa attività deve necessariamente iniziare in corrispondenza del rilascio della prima versione e deve essere ripetuto dal principio ogni volta che una nuova versione viene rilasciata. Al fine di garantire la massima qualità, tutti i membri del gruppo (committente, analisti e project manager) sono coinvolti a vario titolo nella fase di testing e verifica. Criticità possono verificarsi in corrispondenza dell'aumento della velocità di rilascio delle nuove versioni in quanto questa fase impegna tutto il team a lunghe sessioni di lavoro che devono necessariamente essere coordinate.

3.2 Stato dell'arte e possibili sviluppi

La ricerca attualmente è fatta sulle notizie della stampa nazionale ed internazionale passando attraverso aggregatori consolidati che garantiscono un impareggiabile rapporto costo - prestazioni richiesto in progetti con scarsità di risorse economiche.

Il collegamento con le fonti di riferimento (Figura 6) 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 ed i motori di ricerca. L'integrazione di queste informazioni con Twitter e YouTube permette di raccogliere e predisporre un numero 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 IntelligenceIn 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 valutazioni per le fonti e per le informazioni è un elemento necessario per l'aumento della qualità ed attendibilità delle informazioni rilasciate piattaforma stessa.

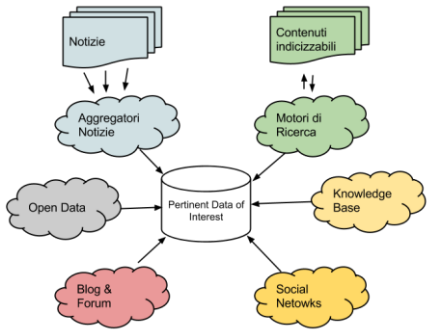

Figura 6 - Dati Pertinenti Analisi

Bibliografia

Alliance, I. S. (2008). *The cyber Security Social Contract: Policy Recomandations ffor the Obama Administration and 111th Congress*. Retrieved 09 2013, from http://www.whitehouse.gov/files/documents/cyber/ISA%20-%20The%20Cyber%20Security%20Social%20Contract.pdf

America, C. (n.d.). Servizi di Informazione e intelligence economica a sostegno della competizione industriale.

autores, L. (2013). BIG DATA: CHALLENGES AND OPPORTUNITIES. Barcelona: Huygens Editorial.

Bookity, S. C. (2015). *Symfony The.* Sesio Labs.

Carlisle, D. (2010, April). *graphicx: Enhanced support for graphics.* Retrieved from http://www.ctan.org/tex-archive/ help/Catalogue/entries/graphicx.html

CLUSIT. (2012 - 2013). *Rapporto CLUSIT*. Retrieved 2013, from CLUSIT: http://www.clusit.it/rapportoclusit/

Colazzo, G. (2014). OSINT: tecniche investigative basate sulle fonti aperte su Internet. Milano: DFA Open Day Università degli Studi Milano.

Confindustria. (2011). *Servizi e infrastrutture per l'innovazione digitale del Paese*. (Confindustria) Retrieved from http://www.fndi.it/download/Rapporto%20servizi%20e%20infrastrutture_executive%20summary.pdf

COPASIR. (2010, Luglio 7). *Comitato di Sicurezza*. Retrieved 09 2013, from Parlamento Italiano: http://www.parlamento.it/documenti/repository/commissioni/bicamerali/COMITATO%20SICUREZZA/Doc_XXXIV_n_4.pdf

Cornford, M. S. (2011). *Total cost of ownership of open source software: a report for the UK Cabinet Office supported by OpenForum Europe.* Queen's Printer and Controller from HMSO.

Jones, S. (n.d.). *Una minaccia persistente.* Retrieved from sicurezzanazionale: http://www.sicurezzanazionale.gov.it/sisr.nsf/letture/una-minaccia-persistente.html

Luciano Floridi, M. (2012). *La rivoluzine dell'informazione.* Codice Edizioni.

Mathieu, L. (2013). *Using OSINT in Your Business.* Retrieved from Infosec Institute: http://resources.infosecinstitute.com/using-osint-in-your-business/

MATTHIAS DEHMER, S. C. (2012). *Statistica and Machine Learning Approaches for Network Analysis.* John Wiley & Sons, Inc.

MongoDB. (2014). *Big Data.* Retrieved from Examples and Gidelines for the Enterprise Decision Makers.

Nacci, G. (2014). Open Source Intelligence Abstration Layer.

Nathan Danneman, R. H. (2014). *Social Media Mining with R.* Packt Publishing.

Ralston, B. (2011). *powerpivot for business intelligence using excel and sharepoint.* Apress*.

Saccone, U. (2014). *Governare i rischi del XXI secolo.* Retrieved from sicurezzanazionale: http://www.sicurezzanazionale.gov.it/sisr.nsf/letture/governare-i-rischi-del-xxi-secolo.html

Steven Roman, P. (2002). *Writing Excel Macros with VBA, 2nd Edition.* O'Reilly.

The Cyber Security Social Contract Policy Recommendations for the Obama Administration. (2013, 09). Retrieved from White House: http://www.whitehouse.gov/assets/documents/Cyberspace_Policy_Review_final.pdf

Voronkov, A. (2004). *EasyChair conference system*. Retrieved from easychair.org

4 Appendice: demo

Di seguito è illustrato il percorso che potrà essere visionato. Nel paragrafo “Funzionalità” sono elencate le principali caratteristiche disponibili nella dimostrazione, mentre nel paragrafo “Casi d’uso” sono illustrate le azioni ed i comportamenti attesi da parte del sistema.

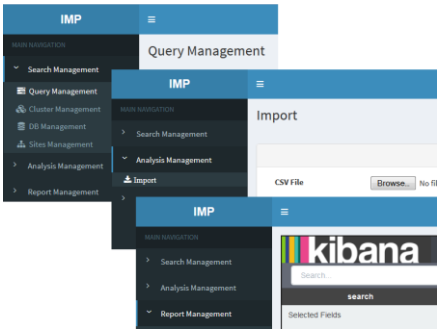

Figura 7 – IMP menu

4.1 Funzionalità

Tramite l'interfaccia applicativa IMP è di tipo web (Bookity, 2015) ed è possibile accedere ai tre moduli principali della piattaforma stessa. Essa si compone del modulo di “Search Management”, da quello di “Analysis Management” e infine da quello di “Report Management”. Tutti e tre i moduli sono facilmente accessibili dal menu di navigazione volutamente molto semplice e lineare ispirato come già detto al ciclo di intelligence classico. E’ possibile prenderne visione in Figura 7.

4.2 Casi d’uso

1. **Organizzazione e gestione dei risultati delle ricerche in contenitori logici**

Questo primo caso d’uso permette di apprezzare il risultato raggiunto e l’operatività della piattaforma. Il primo passo di risposta del processo di Intelligence prevede la pianificazione della ricerca delle informazioni. Per eseguirlo nella Piattaforma è possibile accedere nella prima sezione del menù attualmente indicata con il nome “Search Management”. l’aggregazione dei risultati di ricerca avviene mediante raggruppamenti in contenitori logici: i “Cluster” per definire il più alto livello di aggregazione; “DB” per impostare un livello intermedio; l’ultimo “Query” permette di impostare ed eseguire una query di ricerca. Per ciascuno dei contenitori logici è possibile esportare in forma aggregata tutti i risultati delle queries di ricerca in essi contenute. L’obiettivo principale dopo aver impostato il “contenitore logico” (Figura 8) all’interno del quale inserire i risultati è quello di eseguire una ricerca. Attualmente per una limitazione a livello grafico non è possibile visualizzare la medesima query in più contenitori logici, tuttavia a livello architettonico è stata opportunamente prevista una relazione molti-a-molti per poterlo fare in futuro.

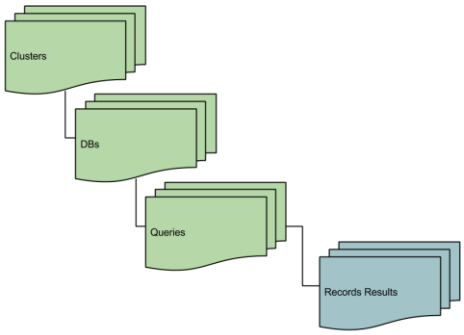

Figura 8 - Clusters, databases, queries e risultati

2. **Esecuzione ricerca sulle fonti OSINT**

Una volta compilata la form e confermata la ricerca mediante l’apposito pulsante di comando “Submit”, presente nella schermata “Query Form” il sistema prende in carico la richiesta ed a seconda dei parametri di ricerca imposta la query verrà eseguita immediatamente oppure verrà posta in stato di “Pending” sino a quando l’orario di esecuzione della stessa sarà sopraggiunto. In particolare è utile far notare che la ricerca su tutte le fonti connesse alla piattaforma è eseguita in parallelo rendendo il processo di ricerca estremamente più rapido ed efficiente.

La ricerca può essere impostata mediante una query form opportunamente strutturata e realizzata, (a tali riguardo vedasi Figura 9). Nella form è possibile specificare diverse impostazioni di ricerca. Per ogni ricerca effettuata è possibile accedere alla sezione di gestione dei risultati ottenuti chiamata appunto "Result Management". Qui è possibile navigare in un opportuno sott'insieme di risultati per poterli consultare e/o poi vedere nel dettaglio. Vi sono immediatamente disponibili le principali informazioni che ne caratterizzano il singolo risultato con il riferimento al link della notizia orginale, nocché la fonte stessa da cui la notizia è stata tratta.

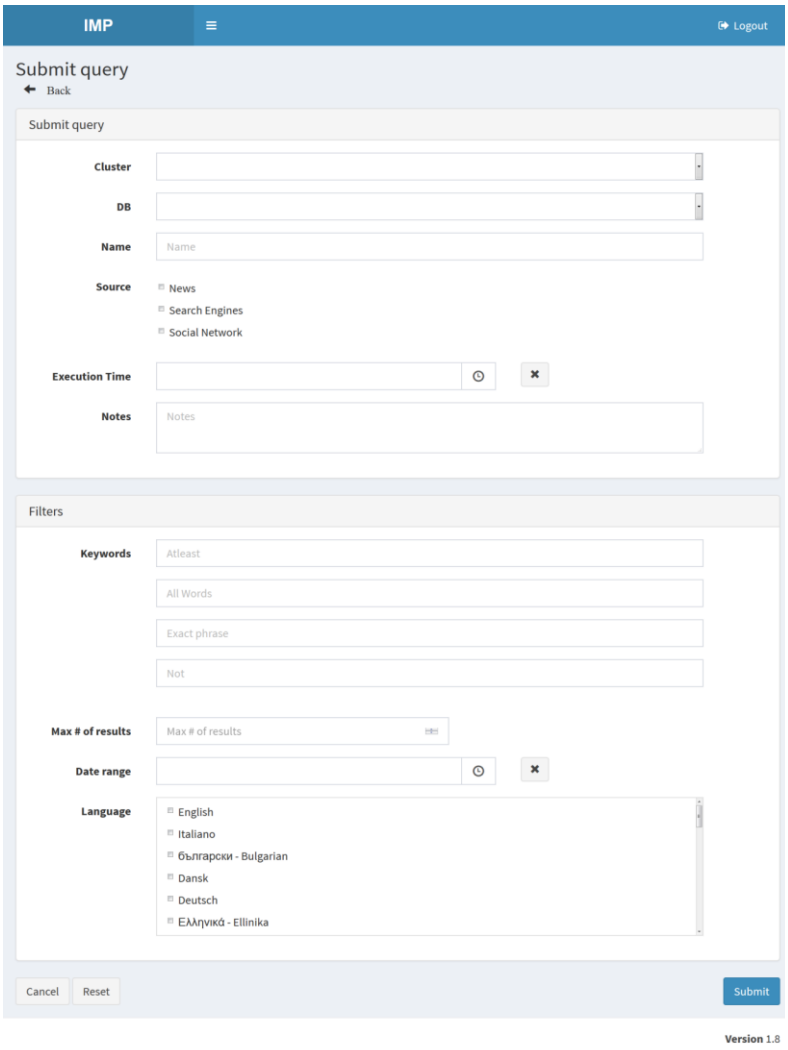

Figura 9 – IMP Query Form

3. Esportazione risultati di ricerca

Per la gestione delle ricerche fatte vi è la sezione (Figura 10) "Query Management". Questa permette di gestire tutte le ricerche effettuate ed avere una rapida interazione tra le stesse mediante l'opportuno utilizzo dei filtri di ricerca con i quali poter ritrovare la Query di interesse. Dalla schermata è possibile passare ai risultati della Query cliccando sul numero di risultati che la query ha generato. In corrispondenza di ogni singola query vi è la funzione "Esporta CSV" che permette la creazione automatica di un file CSV (Comma Separate Value) da utilizzare per effettuare il processo di analisi, utilizzando a sua volta lo strumento ritenuto più adeguato/preferito allo scopo.

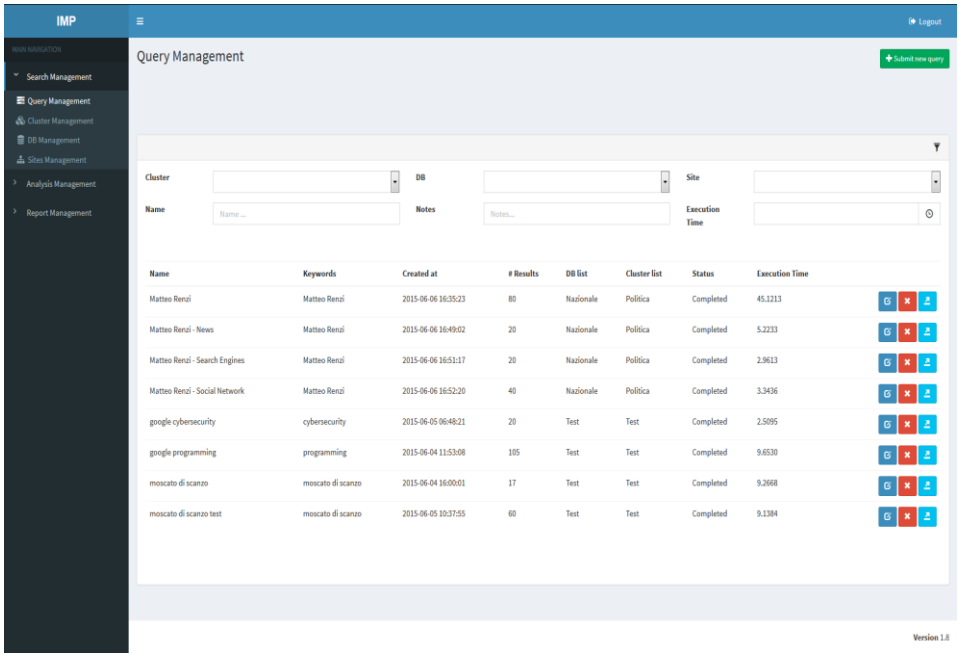

Figura 10 - IMP: Query Management

4. Simulazione processo di analisi

Nella seconda sezione "Analysis Management" è possibile accedere allo strumento oggi disponibile per importare i risultati dell'analisi eseguita sui dati ottenuti dalle ricerche. Come già anticipato in precedenza i risultati delle ricerche sono esportati in un file dati in formato CSV. Una volta esportati i dati questi sono analizzati con un opportuno strumento (per esempio la suite Pentaho oppure gli applicativi Weka e R). Completato il processo di analisi è possibile eseguire l'operazione di importazione di tutti i records precedentemente esportati e poi analizzati ed aggiornati dalle operazioni del processo di analisi. Una volta terminato il processo di importazione sarà possibile visualizzare le informazioni utilizzando il modulo "Report Management". Non essendo disponibile in fase di demo il modulo di analisi verrà simulato importando un file precedentemente opportunamente preparato.

5. Visualizzazione Report informativi

La sezione di "Report Management" è quella nella quale è possibile vedere la visualizzazione dei dati ed è anche possibile la realizzazione di diverse dashboard interattive personalizzate. Le dashboards sono realizzate anch'esse partendo da un progetto Open Source Kibana i cui dettagli tecnologici sono facilmente reperibili sul sito di riferimento liberamente disponibile in Internet. I dati visualizzati sono presi direttamente dal motore di indicizzazione, che a sua volta, grazie alla componente di integrazione "Mongo Connector" li prende praticamente in real-time dal database IMP MongoDB.

Questa stretta integrazione tra database informativo dell'applicativo e cluster indicizzato permette il pieno sfruttamento dello strumento di reportistica visuale. Intanto è possibile avere una idea facendo riferimento allo screen-shoot riportato in Figura 11.

Nella figura sono state preventivamente predisposte due diverse dashboard ciascuna delle quali rappresentante una particolare vista di interesse dei dati. Ogni dashboard permette all'utente di poter visualizzare determinati parametri di interesse, variare a proprio piacimento i valori dell'intervallo analizzato. L'interazione diretta tra analista e dati mediante la visualizzazione di opportuni grafici permette allo stesso di poter individuare eventi significativi e/o di interesse. Pre-impostando

opportunamente le viste e le metriche presenti dashboard riassuntive è possibile in ogni istante avere un immediato riscontro visuale delle ricerche effettuate con il modulo di “Search Management”.

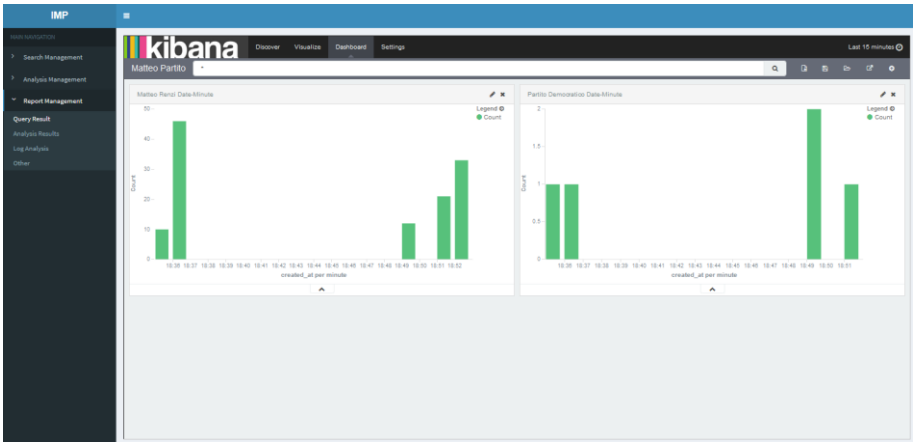

Figura 11 - IMP Report Management