

PEC Qualità Prodotti

Da: hexcogito [hexcogito@pec.aruba.it]
Inviato: venerdì 7 ottobre 2016 18:46
A: qualita.prodotti@aifa.mailcert.it
Cc: francesco.pedini@hexcogito.it
Oggetto: Integrazioni Manifestazione di interesse ai fini dell'eventuale affidamento diretto avente ad oggetto un servizio di IT Intelligence con l'obiettivo di identificare la promozione e l'eventuale offerta di farmaci attraverso i Social Network
Allegati: XC161007_01 - AIFA - Integrazioni Servizio IT Intelligence.pdf

Spett.le Agenzia Italiana del Farmaco,
c.a. Dott.ssa Marta Gramazio,

Con riferimento alla "Manifestazione di interesse ai fini dell'eventuale affidamento diretto avente ad oggetto un servizio di IT Intelligence con l'obiettivo di identificare la promozione e l'eventuale offerta di farmaci attraverso i Social Network" e alla vs. richiesta di chiarimenti/integrazioni ricevuta in data , la HEXCOGITO S.r.l., nella persona del suo Legale Rappresentante, Francesco Pedini, con la presente invia il documento in allegato contenente il dettaglio delle informazioni richieste.

Si rinnova altresì, la propria disponibilità ad eventuali chiarimenti/integrazioni di natura tecnica e/o amministrativa.

Un cordiale saluto

HEXCOGITO S.r.l.

Hexcogito S.r.l.
via Mosca, 52 - 00142 Roma
P.Iva / C.Fisc. 09529541006
amministrazione@hexcogito.it



Roma, 07 ottobre 2016

Spett.le
Agenzia Italiana del Farmaco
Ufficio Qualità dei Prodotti e Contrasto al
Crimine Farmaceutico
Via del Tritone, 169 – Roma
c.a. Dott.ssa Marta Gramazio
qualita.prodotti@aifa.mailcert.it

Integrazioni Servizio IT Intelligence AIFA

Spett.le Agenzia,

Con riferimento alla ns. offerta con data 2 settembre 2016, in atti AIFA prot. n 0089578/A, relativa all'indagine di mercato per l'acquisizione della manifestazione di interesse ai fini dell'eventuale affidamento diretto avente ad oggetto un servizio di IT Intelligence con l'obiettivo di identificare la promozione e l'eventuale offerta di farmaci attraverso i social network, e alla vs. successiva richiesta di chiarimenti con data 4 ottobre 2016, la **HEXCOGITO** riporta nel seguito le integrazioni e i chiarimenti richiesti.

Protocollo offerta:	XC161007_01 del 7 ottobre 2016
Versione del documento	1.0
Autore del documento	Francesco Pedini
Autore della revisione	Francesco Pedini

Cordiali saluti

HEXCOGITO S.r.l.
HEXCOGITO
Via Mosca, 52 - 00142 Roma
P.IVA 09529541006



Indice dei contenuti

1. PRESENTAZIONE DELLA HEXCOGITO S.r.l.	3
1.1 Referenze su attività di IT Intelligence	4
2. EXECUTIVE SUMMARY	5
2.1 Tecnologia utilizzata per l'esecuzione del Servizio.....	5
2.2 Piano di Lavoro	7

1. PRESENTAZIONE DELLA HEXCOGITO S.r.l.

La **HEXCOGITO S.r.l.** opera nel campo dell'Information and Communication Technology (ICT) dal 2007, offrendo prodotti e soluzioni congiunte e all'avanguardia in tutti i settori dell'Informatica, della Formazione e svolgendo attività di analisi, progettazione ed assistenza nella fase di produzione nel campo dell'Ingegneria.

La **HEXCOGITO** vanta un organico composto da professionisti, in gran parte giovani, costantemente formati sulle nuove tecnologie, sugli strumenti di progettazione più richiesti e diffusi nel campo ingegneristico e sulla conoscenza dei prodotti di sviluppo informatico e sistemistico, che consentono di rispondere alle diverse esigenze dei propri Clienti in termini di grande flessibilità e produttività.

Tra le principali attività svolte dalla **HEXCOGITO** figurano contratti strategici per l'attività di produzione nel campo dei servizi ICT, quali ad esempio: *VATICANO, VICARIATO DI ROMA, CONI SERVIZI, SOGIN, LEPIDA, ITALCERTIFER, ISTITUTO SUPERIORE SANITÀ, RISORSE PER ROMA, FATEBENEFRAPELLI ISOLA TIBERINA, C.A.R.A., TECNOLOGIA & SICUREZZA, FEDERSICUREZZA.*

La **HEXCOGITO** è titolare di Software sviluppati internamente ed integra le proprie soluzioni tecnologiche con Software standard di mercato. La **HEXCOGITO** ha un reparto di progettazione e sviluppo Software su piattaforme Windows, Unix/Linux, OSX, per la realizzazione di soluzioni e servizi in ambiente Stand-Alone, Web-Based, Mobile e Multimedia, per sviluppo di sistemi di pianificazione e controllo di gestione in ambito Business Solution e Business Integration, di sistemi di Content Management, Document Management, di riconoscimento e trattamento di immagini, di riconoscimento e trattamento vocale.

La **HEXCOGITO** eroga servizi di Business Intelligence attraverso l'analisi delle esigenze, a modellazione di processi e data model, ed estrazione dei dati. Creazione di basi di dati affidabili, certificate e consistenti. Realizzazione ed organizzazione dei processi di alimentazione con il minimo impatto architeturale per i sistemi sorgenti. Realizzazione ed assistenza nell'utilizzo del sistema di visualizzazione e reporting. Realizzazione di sistemi di controllo e cruscotti aziendali e di funzioni di pianificazione e predittive.

La **HEXCOGITO** eroga servizi per la revisione, automazione e ottimizzazione dei flussi informativi e dei processi, insieme alla progettazione e alla realizzazione di applicazioni informatiche per la gestione integrata delle informazioni e delle conoscenze aziendali.

La **HEXCOGITO** eroga servizi per la gestione di Infrastrutture e Sistemi, attività di progettazione e realizzazione di infrastrutture IT in ambienti tradizionali e virtualizzati, LAN, WAN, Cloud, assistenza sistemistica su sistemi UNIX, Microsoft, Oracle, è in grado di gestire internamente tutte le attività sistemistiche e di configurazione network, con conseguente abbattimento dei costi legati alle manutenzioni ordinarie e straordinarie.

La **HEXCOGITO** opera dal 2016 in regime di Qualità Controllata e ha ottenuto la certificazione del Sistema di Qualità, conforme alla normativa **ISO 9001:2008** per i settori **IAF 19, 33, 35, 37** (n. certificato **HEXC5854Q1601**) e la **ISO/IEC 27001:2013** per i settori **IAF 33, 35, 37** (n. certificato **HEXC5854O1601**), rilasciati dall'istituto EuCI (European Certification Institute)

1.1 **Referenze su attività di IT Intelligence**

[eventuali riferimenti e/o referenze su precedenti attività svolte rispetto a ricerche di IT intelligence su social network, con particolare riguardo al settore farmaceutico e alla vendita di farmaci attraverso il web]

- **Tecnologia & Sicurezza:** servizio di IT Intelligence per la ricerca e l'estrazione di dati relativi a dispositivi di protezione individuale (DPI) nel campo della sicurezza sui luoghi di lavoro. L'attività è stata condotta attraverso l'esecuzione di campagne di analisi bimestrali sui Social Network supportati dal Framework della **HEXCOGITO** nel corso degli anni 2014, 2015 e 2016.
- **Risorse per Roma:** servizio di IT intelligence per la ricerca e l'estrazione di dati relativi ad abusi edilizi. L'attività è stata condotta attraverso l'esecuzione di una campagna sui Social Network supportati dal Framework della **HEXCOGITO** nel corso dell'anno 2013.
- **ENAV:** analisi, progettazione e sviluppo di una piattaforma di IT intelligence per la ricerca e l'estrazione di dati relativi al volo su Web e Social Network. L'attività è stata eseguita nel corso dell'anno 2013.
- **Istituto Superiore Sanità:** servizio di IT intelligence per la ricerca e l'estrazione di dati relativi a particolari principi attivi. L'attività è stata condotta attraverso l'esecuzione di una campagna semestrale sui Social Network supportati dal Framework della **HEXCOGITO** nel corso dell'anno 2011.

2. EXECUTIVE SUMMARY

L'avvento dei Social Network ha determinato la generazione e la memorizzazione di una quantità enorme di informazioni. I Social Network rappresentano quindi una fonte potenzialmente infinita (e in alcuni casi pericolosa) di dati degli utenti, i quali possono essere sfruttati con diversi scopi, primo fra tutti quello commerciale.

L'estrazione dei dati dai Social Network è attualmente uno degli argomenti di maggior interesse. Infatti, tali informazioni offrono molte opportunità di ricerca e analisi: analizzando i dati di una rete sociale è possibile determinare una rappresentazione approssimata del mondo reale.

L'obiettivo di questo capitolo è quello di illustrare l'approccio metodologico utilizzato da **HEXCOGITO** per gestire, anche dal punto di vista organizzativo, le attività e l'erogazione del servizio IT intelligence per identificare la promozione e l'offerta di farmaci attraverso i social network (SN).

Le modalità di svolgimento del servizio recepiscono le esperienze maturate da **HEXCOGITO** nel corso di analoghe forniture di servizi per altri contesti pubblici e privati. Le modalità operative sono, inoltre, coerenti con le indicazioni fornite dal CNIPA. Nell'ambito del servizio è, peraltro, **garantita la soddisfazione delle esigenze estemporanee, non pianificabili, o insorte a causa di eventi imprevisti cui si deve dare risposta pressoché immediata.**

2.1 Tecnologia utilizzata per l'esecuzione del Servizio

[Tipo di tecnologia/software/piattaforma che si intende utilizzare ai fini dell'esecuzione del servizio richiesto]

I Social Network, al fine di permettere lo sviluppo di applicazioni attorno alle proprie piattaforme, mettono a disposizione delle interfacce (*Application Programming Interface o API*) per permettere l'accesso, nel rispetto della normativa legata alla privacy, ai dati dei propri utenti.

Tuttavia, considerato che i Social Network poggiano il proprio modello di business proprio su tali dati, ne consegue che le modalità e la quantità dei dati accessibili attraverso le API è estremamente ridotta. Conseguentemente è necessario prendere in considerazione altre tecniche che permettano la ricerca e l'estrazione di informazioni dai Social Network in modo efficace.

La **HEXCOGITO**, attraverso il proprio Team di Ingegneri Informatici, ha maturato una specifica esperienza in campo accademico, grazie alla pluriennale collaborazione con alcune Università romane, impegnate in attività di Ricerca di nuovi modelli semantici e di prototipi per la ricerca e l'estrazione di dati dal Web e dai Social Network.

La **HEXCOGITO** ha successivamente avviato un progetto di Ricerca & Sviluppo interno finalizzato alla realizzazione di un *Framework proprietario di IT Intelligence* per la ricerca, l'analisi, il processamento e l'estrazione di dati dal Web e dai Social Network.

Il Framework è stato sviluppato con tecnologia Open Source e realizzato attraverso l'adozione di un modello di sviluppo S/W basato in parte sul modello a cascata (*waterfall model*), una metodologia di progettazione sub-sequenziale per lo sviluppo S/W, concepita in più fasi correlate, e influenzato massivamente da metodi adattivi dello sviluppo Agile (*agile development methods*), in cui i requisiti e le soluzioni evolvono attraverso un approccio iterativo ed incrementale, incoraggiando risposte flessibili ai cambiamenti richiesti dal Web.

Il motore implementato all'interno del Framework integra le API rese disponibili dai Social Network unitamente a delle tecniche di estrazione automatica, comunemente definite *Web Crawling and Scraping*. Il Web Scraping si riferisce ad un insieme di tecniche per l'estrazione automatica dei contenuti dai siti Web via HTTP, finalizzate alla trasformazione di tali contenuti in informazioni strutturate e quindi utilizzabili. In altre parole, il processo di Web Scraping può essere descritto come il processo di estrazione dei dati da pagine HTML, conseguentemente qualsiasi contenuto Web visibile all'utente attraverso un Browser può essere scaricato e processato.

Le fasi principali in cui si compone l'intero processo sono le seguenti:

1. Ricerca e recupero delle pagine HTML (i. e. *crawling*)
 - a. Gestione della paginazione
 - b. Gestione delle richieste asincrone
2. Estrazione e strutturazione dei dati (i.e. *scraping*)
3. Export

I processi di *Crawling and Scraping*, utilizzati in senso assoluto, potrebbero violare i termini di contratto e di utilizzo dei siti target, a seconda delle finalità di utilizzo dei dati estratti. In generale, i termini sono specificati all'interno del file *robots.txt* e il Framework implementa un algoritmo per la verifica e la gestione degli accessi dipendentemente dalle politiche adottate dalle specifiche piattaforme Social in totale sicurezza.

Il Framework è compatibile con i seguenti Social Network:

- Facebook
- Flickr
- Google+
- Instagram
- LinkedIn
- Reddit
- Tumblr
- Twitter
- Youtube
- Vimeo

Dal punto di vista funzionale, il Framework mette a disposizione un'interfaccia dove è possibile inserire i criteri di ricerca e specificare i Social Network di interesse. È possibile impostare delle ricerche personalizzate, attraverso l'inserimento di chiavi di ricerca multiple (i.e. frasi esatte, keyword, hyperlink, hashtag, etc...) correlate da diversi operatori logici (AND, OR, NOT), permettendo l'esecuzione di ricerche complesse all'interno dei Social Network.

Il Framework è progettato seguendo lo standard *ISO 18N* e supporta attualmente le seguenti lingue: *italiano, inglese, francese, tedesco, spagnolo, portoghese*. Sono invece in corso di integrazione le seguenti lingue: *brasiliiano* e il *polacco*.

Il Framework fornisce un meccanismo di aggiornamento in tempo reale di insiemi di informazioni definite relative ad oggetti e connessioni. Questa funzionalità permette, una volta estratto un insieme di dati di interesse, di monitorare l'evoluzione degli stessi.

Il Framework permette di importare la lista delle parole chiavi attraverso il caricamento di un file CSV. È possibile inoltre, consultare i dati estratti direttamente online e prevede la possibilità di esportare i dati in diversi formati (XLS, CSV, XML, etc...).

Il Framework sviluppato dalla **HEXCOGITO** allo stato attuale non implementa in modo nativo le funzionalità per eseguire ricerche sui seguenti Social Network: *Blogger* e *Pinterest*. La **HEXCOGITO**, con l'obiettivo di estendere la ricerca a TUTTI i Social Network richiesti dall'Agenzia, intende utilizzare le funzioni di ricerca messe a disposizione direttamente dai Social Network e motori di ricerca sviluppati da Terze Parti. I dati ricevuti saranno quindi integrati ed armonizzati con l'Output Dataset per le successive attività di post-processamento previste nella Fase 2 e nelle successive.

2.2 Piano di Lavoro

[Piano di lavoro operativo e relative tempistiche, tenendo presente comunque la scadenza fissata al 31 ottobre 2016 per la consegna dei risultati dell'indagine (max 8.000 battute spazi inclusi)]

Al fine di garantire all'Agenzia un adeguato livello di servizio, la necessaria competenza tecnica e la qualità necessaria per un appalto pubblico, di seguito esponiamo nel dettaglio le **modalità** e le **metodologie** che saranno utilizzate da **HEXCOGITO** per dare seguito alle attività richieste.

Nel seguito è illustrato il *Piano di Lavoro* proposto dalla **HEXCOGITO** per l'erogazione del servizio di IT intelligence per identificare la promozione e l'offerta di farmaci attraverso i social network (SN).

Importante: Precisiamo sin da ora che, nel caso in cui taluni elementi qualificanti della nostra metodologia organizzativa dovessero inavvertitamente risultare non completamente aderenti al modello atteso, sarà sempre salvaguardata la condizione di maggior favore per l'Agenzia.

Poiché le diverse fasi che compongono il servizio richiesto dall'Agenzia si caratterizzano per una forte interrelazione (nella misura in cui spesso i risultati di una o più di esse servono da input ad una o anche più di un'altra, giungendo a condizionarne il risultato), **HEXCOGITO** ritiene

fondamentale adottare un modello di gestione più che collaudato, basandosi su standard ben precisi (ISO 9001:2008, ITIL v.3), in modo da assicurare la ripetibilità del successo già ottenuto in altre situazioni.

Il Responsabile di Progetto nominato dalla **HEXCOGITO** sarà costantemente il punto di riferimento per l'erogazione del servizio nei confronti dell'Agenzia e seguirà il progetto sia in avviamento che a regime, facendosi garante della qualità delle operazioni e di eventuali azioni migliorative ai fini del raggiungimento degli obiettivi.

Sarà cura del Responsabile di Progetto **HEXCOGITO** gestire efficacemente l'attività definita nel Piano di Lavoro e le richieste, che possono giungere secondo tempistiche e volumi anche non predicibili. In tal senso, le strutture e le metodologie di **HEXCOGITO** garantiscono la corretta ed efficace ricezione delle richieste e la pronta reazione in termini di allocazioni delle risorse.

Il Responsabile di Progetto **HEXCOGITO** si occuperà di:

- assegnare le risorse disponibili alle attività da espletare;
- controllare la corretta ed efficace erogazione delle attività, vagliando i risultati e la qualità;
- garantire la produzione della documentazione corrispondente e il rilascio all'Agenzia.

Il Piano di Lavoro del Servizio sarà organizzato nelle seguenti macro-fasi:

FASE	DESCRIZIONE
FASE 1	Monitoraggio dei Social Network
FASE 2	Ricerca ed Estrazione
FASE 3	Identificazione dei collegamenti
FASE 4	Feedback
FASE 5	Consegna

Nella "FASE 1 – MONITORAGGIO DEI SOCIAL NETWORK" sarà preliminarmente definito l'*Input Dataset*, costituito da 90 voci complessive in 4 lingue (italiano, inglese, spagnolo e portoghese), da utilizzare durante il monitoraggio, la ricerca e l'estrazione dei dati dai Social Network. Il Dataset di input sarà caricato nel Framework della **HEXCOGITO**, congiuntamente alla configurazione di utilizzo.

Nella "FASE 2 – RICERCA ED ESTRAZIONE" saranno eseguite le sessioni di ricerca degli elementi appartenenti all'Input Dataset. Per ogni entry trovata (pagina SN e/o farmacia online), saranno memorizzati e categorizzati i seguenti elementi:

- screenshot;
- nome del dominio;
- paese che effettua l'hosting;
- classificazione della pagina web;
- lingua del sito/pagina di SN;
- top level domain;
- il nome del prodotto citato;
- la presenza di espressione "senza prescrizione";
- hyperlink (URL/dominio/nome del dominio);
- URL/dominio/nome del dominio delle pagine di re-indirizzamento.

Con l'obiettivo di ottimizzare i tempi di progetto, le sessioni di ricerca saranno eseguite in parallelo sui diversi Social Network, sfruttando le potenzialità dell'Infrastruttura Cloud a disposizione della **HEXCOGITO**.

Compatibilmente con i requisiti di Progetto e la scadenza definita dall'Agenzia (31 ottobre 2016), il Framework della **HEXCOGITO** provvederà ad eseguire nuove esecuzioni, sfruttando le funzionalità di aggiornamento in tempo reale sull'insieme dei dati estratti all'esecuzione precedente, con l'obiettivo di monitorare l'evoluzione degli stessi.

Una volta disponibili i risultati (*Output Dataset*), sarà eseguita una scansione per individuare le farmacie online. Sui record relativi alle farmacie online saranno eseguite delle attività di post-processamento, atte a verificare l'appartenenza alle liste autorizzate e ad estrarre ulteriori dati relativi ai contatti della farmacia. I dati individuati saranno integrati nell'Output Dataset.

Nella "FASE 3 – IDENTIFICAZIONE DEI COLLEGAMENTI" saranno eseguite delle attività atte ad individuare dei collegamenti tra le diverse pagine dei SN e tra le diverse farmacie online individuate, sulla base di criteri di associazione legati a indirizzi IP, Hyperlink e domini comuni.

Nella "FASE 4 – FEEDBACK" sarà eseguito il controllo atto a verificare che il numero dei falsi positivi (tollerabili) e dei falsi negativi (meno tollerabile) sia accettabile. I risultati della Fase 4 daranno luogo a delle conseguenti correzioni sul processo di monitoraggio, ricerca ed estrazione.

Nella "FASE 5 - CONSEGNA" si completerà l'attività con la consegna dei risultati conseguiti unitamente ad una relazione dettagliata delle attività eseguite nelle lingue italiano ed inglese. I dati estratti dalla piattaforma saranno esportati in formato Excel (se necessario anche in formato CSV e/o XML) e resi disponibili all'Agenzia.

La HEXCOGITO, fissando il Kick Off (TO) dell'attività al **17 ottobre 2016**, nel rispetto della scadenza improrogabile indicata dall'Agenzia (**31 ottobre 2016**), intende proporre il seguente *Schedule*:

FASE	DESCRIZIONE	SCADENZA
FASE 1	Monitoraggio dei Social Network	$T1 = TO + 2$
FASE 2	Ricerca ed Estrazione	$T2 = T1 + 5 (*)$
FASE 3	Identificazione dei collegamenti	$T3 = T2 + 2$
FASE 4	Feedback	$T4 = T3 + 2$
FASE 5	Consegna	$T5 = T4 + 1$

(*) La Fase 2 sarà eseguita continuativamente sia nei giorni feriali che festivi (compresi sabato e domenica).

Importante: Precisiamo sin da ora che, nel caso fosse possibile anticipare la data di Kick Off e/o l'Agenzia decidesse di posticipare la data di consegna, HEXCOGITO si rende disponibile ad ampliare il periodo di osservazione delle Fasi 2-4, sfruttando le funzionalità di aggiornamento in tempo reale dell'applicazione per svolgere ulteriori analisi, ricerche ed estrazione dei dati, con l'obiettivo di aggiornare l'Output Dataset.

HEXCOGITO S.r.l.
HEXCOGITO
Via Mosca, 52 - 00142 Roma
P.IVA 09529541006