Estrazione dei dati sensibili da Facebook e Linkedin

Nicola Valentino Messuti  
Università degli studi di Salernon.messuti2@studenti.unisa.it

*Abstract*— I social network hanno assunto un enorme importanza nel mondo moderno e la quantità di dati sensibili presente su queste piattaforme ha creato nuove problematiche legate alla privacy di tali informazioni. Numerosi studi hanno analizzato ed osservato i rischi associati con la condivisione dei nostri dati sui vari social media che possono portare ad attacchi mirati su individui, associazioni ed aziende e, negli anni, si è avuta una significativa attività legislativa volta a mitigare tali rischi. Nonostante gli studi e le leggi più recenti, la problematica della privacy sui social rimane un problema aperto in quanto tali siti nascono con lo specifico scopo di interconnettere gli individui tramite lo scambio di informazioni strettamente personali. Il lavoro svolto si è posto come obbiettivo quello di approfondire la quantità ed il tipo di dati personali liberamente accessibili sui social media e di individuare una soluzione al problema dell’estrazione di tali dati. A tale scopo sono stati presi in analisi due piattaforme social specifiche, Facebook e Linkedin, su cui sono stati estratti i dati personali di un ristretto numero randomico di utenti di tali piattaforme. La soluzione implementata si basa sul combinare l’utilizzo di due strumenti open-source reperiti online, social mapper e profiler, e di uno strumento realizzato ad hoc. I risultati del lavoro sono un database dei dati estratti dagli utenti ed un’analisi statistica sulle informazioni ottenute.

# Introduzione

I social network sono rapidamente diventati uno dei principali mezzi di scambio e diffusione delle informazioni sul web. A partire dal 2021, il numero di persone che utilizzano i social media è di oltre 3,96 miliardi in tutto il mondo (50,6% della popolazione mondiale), con ogni utente che possiede in media 8,8 account social. Oltre il 66,09% di utenti di piattaforme popolari come Facebook, inoltre, utilizzano i social media ogni giorno [1]. Dalle statistiche si può facilmente intuire perché i social siano diventati una parte fondamentale del mondo moderno e perché la comunicazione su tali piattaforme abbia assunto un ruolo particolarmente rilevante in numerosi ambiti.

I social vengono considerati dalla maggior parte degli utenti come strumenti di comunicazione personale e l’importanza di salvaguardare tali informazioni è spesso data per scontata [5]. Con il passare del tempo, gli utenti inseriscono sempre più dati all’interno di tali piattaforme e questo porta ad una crescita progressiva delle informazioni personali liberamente disponibili ad un vasto gruppo di attori web [5]. Questi attori sono di diversa natura: alcuni potrebbero sfruttare tali informazioni per delle pubblicità mirate ed altri per vere e proprie operazioni di phishing e scam [5]. Sorge in ogni caso una problematica legata alla privacy di tali dati che riguarda non solo un concetto di consapevolezza degli utenti ma anche delle piattaforme stesse che devono tenere conto di eventuali leggi e normative degli stati. In Europa la principale normativa sulla privacy è la General Data Protection Regulation (GDPR) [2] in vigore in tutti gli stati membri dal 25 maggio 2018. Uno degli obbiettivi principali del GDPR è stato quello di restituire ai cittadini il controllo dei propri dati personali [2], controllo che era andato a diminuire negli anni con lo sviluppo di nuove piattaforme digitali che avevano ampiamente superato le capacità delle normative vigenti fino a quel momento. Nonostante la regolamentazione, come testimoniano alcuni articoli e ricerche sulla tematica della privacy come il [3], possiamo osservare che il dibattito è ancora particolarmente acceso su tali temi e che si tratta di un campo in continua evoluzione che necessita di studi continui, di nuove metodologie e punti di vista.

# Lavori correlati

In seguito, viene riportato un piccolo estratto di lavori e studi scientifici che riguardano i principali argomenti correlati con il lavoro svolto e che, quindi, trattano le problematiche di privacy ed estrazione dei dati dai social network:

1. *A Critical Analysis of Privacy and Security on Social Media*

In questo paper [3] vengono analizzate alcune delle problematiche relative alla privacy e sicurezza dei dati sui social e vengono fornite alcune relative metodologie risolutive.

1. *Social Network Extraction of Academic Researchers*

Questo paper [4] è incentrato sul problema dell’estrazione e profilazione dei dati dai social networks. L’obbiettivo principale del lavoro è quello di trasformare un problema solitamente risolto ad hoc in un problema formalizzato e risolto per via automatica.

1. *Online social networks security and privacy: comprehensive review and analysis*

Un paper [5] focalizzato sulla privacy dei dati personali degli utenti. Il lavoro si concentra sulle principali problematiche e le relative soluzioni esistenti. Il lavoro contiene anche diverse statistiche relative agli attacchi sulle varie piattaforme.

1. *Privacy preserving social network data publication*

Questo paper [6] presenta un’analisi comprensiva degli sviluppi recenti nell’ambito della privacy, dei rischi e degli attacchi relativi ai dati personali sui social media con un focus specifico sulle metodologie e tecniche di protezione della privacy. Il lavoro riportato ha come obbiettivo quello di aiutare il lettore nella comprensione di tali tematiche e di analizzare possibili sviluppi futuri nell’ambito della privacy.

1. *Privacy Protection for Social Networking Platforms*

I social network moderni integrano contenuti di terze parti all’interno delle proprie piattaforme e forniscono agli sviluppatori di tali contenuti accesso ai dati degli utenti. Il lavoro [7] cerca di risolvere la problematica relativa alla privacy di tali dati implementando delle API basate su un design privacy-by-proxy. Lo studio si è basato su 150 applicazioni di Facebook ed ha osservato che la stragrande maggioranza di tali applicativi può mantenere la propria funzionalità utilizzando una quantità limitata di informazioni fornita tramite un’interfaccia che mostra solamente dati anonimi o fittizi.

1. *Privacy Social Network Sites and Social Relations*

Il lavoro [8] si concentra su un ristretto campione di utenti a cui vengono fatte domande su privacy e sicurezza relative ai social media. Vengono raccolti dati sulle loro opinioni e sugli attacchi ricevuti ed i risultati vengono analizzati e discussi in relazione alle principali problematiche relative alle più popolari piattaforme social.

1. *Social Media Users’ Legal Consciousness About Privacy*

Questo articolo [9] esplora i vari modi in cui il concetto di privacy è concepito nel mondo moderno. L’articolo si focalizza sull’analizzare i modi in cui un campione di utenti di vari social media gestisce le proprie informazioni personali, e quanto il loro livello di coscienza legale relativo alle leggi sulla privacy influisce sulle loro scelte.

1. *Social Media Web Scraping using Social Media Developers API and Regex*

L’informazione condivisa sui social network è spesso non rilevante per l’utente o difficile da accedere poiché ridondante. In questo lavoro [10] viene proposto un sistema basato sul web scraping capace di estrarre dati dai social, elaborarli e presentarli all’utente in una forma migliore. Il sistema sfrutta le API dei siti specifici ed è particolarmente interessante poiché fornisce un metodo per estrapolare informazioni dai vari social network.

1. *Social media analytics: a survey of techniques, tools and platforms*

Il paper [11] fornisce un’analisi complessiva dei principali tools di social networking come social media, blog e wiki. È principalmente un testo analitico, utile poiché fornisce una descrizione dettagliata delle principali piattaforme social.

1. *Social media analytics – Challenges in topic discovery, data collection, and data preparation*

L’analisi delle informazioni presenti sui social media coinvolge 4 differenti fasi: data discovery, data collection, data preparation e data analysis. Se sulla fase di data analysis troviamo un elevata quantità di ricerche e lavori scientifici, lo stesso non si può dire delle altre tre fasi che mancano della stessa attenzione dedicata alla quarta fase. Questo lavoro [12] si propone di ridurre tale gap conducendo un’analisi estesa e strutturata sulle principali sfide e relative soluzioni proposte all’interno di tali fasi. L’articolo è pensato per essere letto ed approfondito da studenti e ricercatori interessati nell’estrarre ed analizzare dati dai social media.

1. *How to address data privacy concerns when using social media data in conservation science*

Questo articolo [13] analizza le basi legali per l’uso dei dati estratti dai social media con particolare attenzione al rispetto del diritto di privacy degli utenti così come definito dalla normativa europea sulla protezione dei dati personali (GDPR) e discute diversi approcci atti al ridurre i rischi associati con l’estrazione, l’uso e la conservazione dei dati personali degli utenti.

# Social network analizzati

Il lavoro svolto si è concentrato sull’analisi di due social network specifici: Facebook e Linkedin. I dati personali conservati da tali piattaforme possono essere suddivisi in 4 principali tipologie:

1. Lista di amici o contatti – contiene una lista di tutti gli utenti di cui l’utente ha accettato la richiesta di amicizia (o di contatto). Dipendentemente dalle impostazioni di privacy di ogni utente, questa informazione può essere visibile solo a tali utenti oppure essere pubblica.
2. Informazioni personali – include tutte le informazioni che l’utente ha deciso di condividere sul suo profilo come interessi personali, professione, età, orientamento politico e/o religioso, appartenenza a gruppi sul social, ecc.
3. Post – messaggi pubblici scritti dall’utente stesso o ricevuti da altri utenti o da applicazioni. Questo tipo di dati solitamente contiene informazioni sullo stato d’umore dell’utente, le attività che sta svolgendo in quel momento o un’interazione con un altro utente.
4. Foto – tutte le immagini caricate dall’utente. L’immagine del profilo è visibile a chiunque, mentre tutte le altre sono limitate dall’impostazione di privacy dell’utente.

Per tutte le tipologie eccetto la 2, le informazioni condivisibili sui due social sono grosso modo equivalenti. Il lavoro si è concentrato sull’analizzare ed estrarre le informazioni personali (punto 2) reperibili pubblicamente sulle due piattaforme. I dati sono stati estratti, quindi, senza possedere un’amicizia o un contatto con i relativi utenti interessati. Per Facebook i dati estraibili sono il lavoro, l’indirizzo di residenza e l’istruzione mentre per Linkedin sono il lavoro, l’indirizzo di residenza, sito web personale, e-mail e data di compleanno. Come si può notare, Facebook mette in chiaro una quantità di informazioni decisamente inferiore rispetto a Linkedin e due di tali dati, ovvero lavoro e indirizzo, sono in comune tra i due social. In generale un analisi cross-social tra queste due piattaforme consente di ottenere un numero maggiore di informazioni: possiamo, ad esempio, estrarre da un ipotetico utente (che possiede un profilo su entrambi i social) le immagini condivise online su Facebook e tutte le informazioni personali presenti su Linkedin.

# Strumenti: social mapper, Profiler, Custom Crawler

Nel corso del progetto sono stati utilizzati due strumenti reperiti online ed uno strumento realizzato ad hoc:

* **Social Mapper[14]**: uno strumento open source di ricerca di profili social che ci permette di individuare un insieme di utenti (su diversi siti) partendo dall’immagine e dal nome del profilo.
* **Profiler[15]**: uno strumento open source che genera un insieme di link a vari social network partendo da una coppia nome-cognome o da un username.
* **Custom Crawler**: lo strumento realizzato ad hoc che va ad estrarre nome, cognome, dati in chiaro ed immagine del profilo partendo da una lista di link di Facebook.

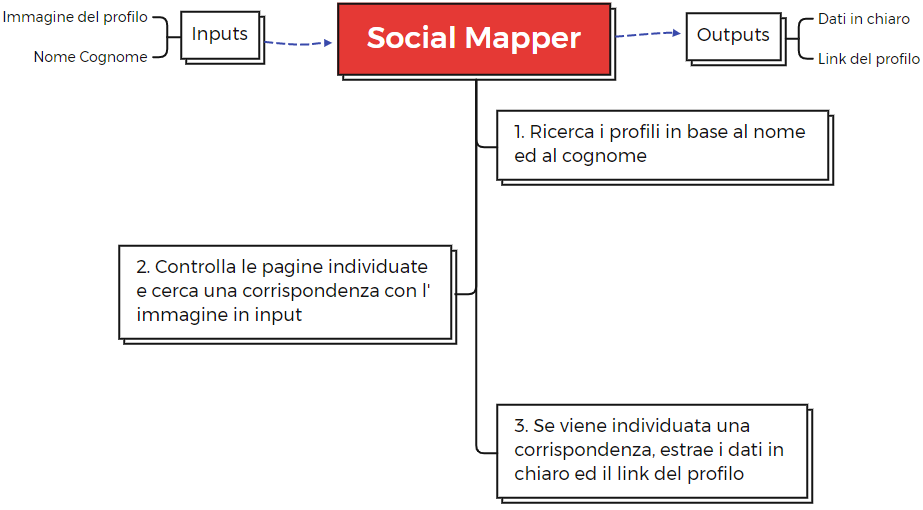


Figura 1: Social mapper

Social mapper è realizzato interamente in Python e sfrutta le librerie BeautifulSoup, Selenium e Dlib. Il funzionamento di social mapper è grosso modo lineare:

1. Effettua il log-in sul sito web interessato
2. Ricerca il profilo in base al nome ed al cognome
3. Controlla i risultati e cerca una corrispondenza con l’immagine di input
4. Se trova una corrispondenza, estrae il link del profilo e gli eventuali dati condivisi dall’utente

La libreria Selenium [16] consente di automatizzare log-in, ricerca ed estrazione del testo html dalle pagine web. Dlib [17] è usato per effettuare un confronto tra l’immagine in input e l’immagine del profilo. BeautifulSoup [18] semplifica il lavoro di estrazione dati da testi in formato html e xml.

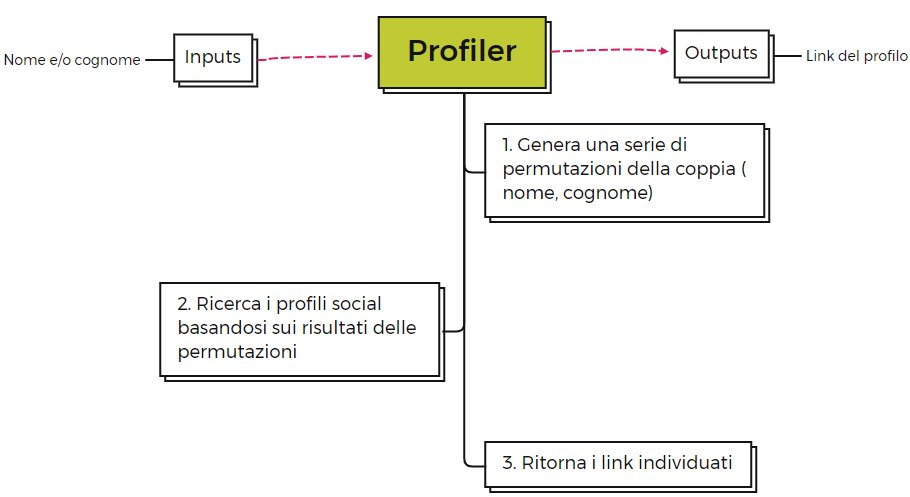


Figura 2: Profiler

Anche profiler è realizzato interamente in Python. A differenza di social mapper, sfrutta le API specifiche di ogni sito per effettuare la ricerca dei profili. Per ogni coppia nome/cognome può potenzialmente individuare più di un link per ogni profilo poiché lavora su una lista di permutazioni della coppia nome/cognome: per ogni elemento della lista controlla se esiste un link al sito che termina con tale valore.

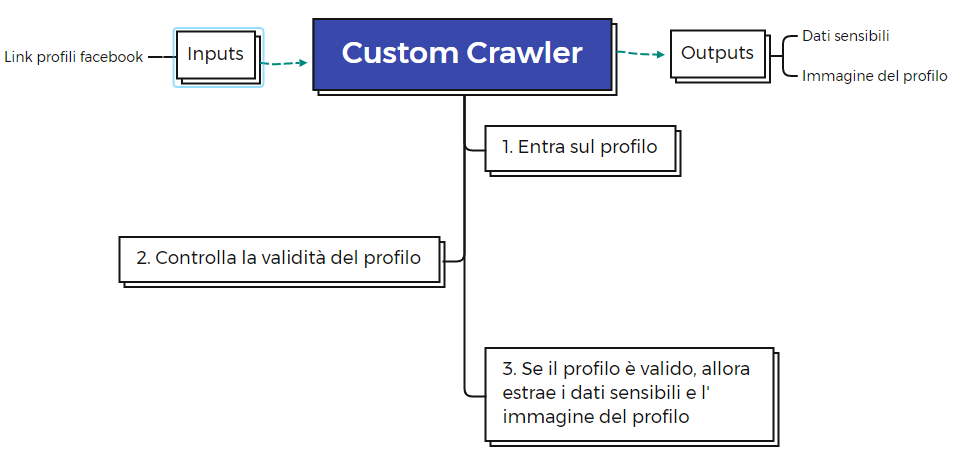


Figura 3: Custom Crawler

Il crawler realizzato ad hoc per Facebook implementa una soluzione simile a quella utilizzata per Social mapper e vengono sfruttate le stesse librerie per gli stessi scopi. La differenza principale è che, conoscendo a priori i link ai profili, possiamo saltare la ricerca degli stessi e limitarci a controllare la validità del link. E’ stato implementato specificatamente per lavorare sull’output di social mapper, che, nel caso del lavoro svolto, si traduce in una lista di link di pagine di potenziali profili Facebook. Il crawler controlla prima di tutto la validità del profilo che sarà considerato come valido se e soltanto se l’utente ha un’immagine del profilo e se su tale immagine viene rilevato solamente un volto. Se il profilo risulta valido, allora il programma procede con l‘estrazione dell’immagine e degli eventuali dati personali.

# Estrazione dei dati

L’estrazione automatica di dati e relative statistiche da Facebook e Linkedin presenta diverse problematiche:

* Un primo problema è la presenza di un numero limitato di strumenti software reperibili online.
* Un secondo problema è l’aggiornamento frequente delle pagine web dei siti social che possono rendere complesso il realizzare dei data crawler funzionanti a lungo termine.
* Un terzo ed ultimo problema è la presenza di sistemi di sicurezza che possono individuare ed interrompere l’uso di processi automatici.

La prima problematica ci obbliga a realizzare degli strumenti ad hoc o a modificare gli strumenti esistenti e adattarli al nostro obbiettivo specifico. Sia i data crawler reperibili online che quelli eventualmente realizzati ad hoc ricadranno però nella seconda problematica e dovremo regolare il nostro lavoro di conseguenza. Nel caso del progetto svolto, l’obbiettivo è stato limitato al realizzare un dataset statico delle informazioni di un insieme di utenti individuato randomicamente da cui fosse possibile estrarre delle statistiche relative ai due social presi in analisi. La terza ed ultima problematica varia in base al social specifico:

* **Linkedin:** ha una struttura html grosso modo stabile, subisce pochi aggiornamenti e tutte le informazioni sensibili di ogni contatto sono mostrate in chiaro su di una pagina separata. In termini di sicurezza, l’accesso alle informazioni sensibili richiede un log-in sul sito. Una ricerca automatica su un elevato numero di contatti (dell’ordine delle centinaia) viene riconosciuta dal sito che forza un log-off dell’account; al log-in successivo, viene richiesto un codice captcha che va ad ostacolare tale ricerca e che forza, quindi, un intervento manuale. Diviene, di conseguenza, complesso realizzare uno strumento completamente automatico per l’estrazione dei dati personali da Linkedin.
* **Facebook:** ha una struttura html considerevolmente più dinamica e soggetta a cambiamenti rispetto a Linkedin. Le informazioni sensibili sono contenute all’interno della pagina principale di un utente e devono essere separate dal resto dei dati. Le informazioni sono mostrate in chiaro anche senza log-in se l’utente non ha impostazioni particolari di privacy. E’ possibile realizzare un sistema completamente automatico per l’estrazione delle informazioni, ma tale sistema richiede maggiore manutenzione rispetto a Linkedin. Degna di una particolare nota è l’uso da parte di Facebook di classi con valori randomici all’interno degli elementi html delle pagine: questa peculiarità può complicare lo sviluppo di un data crawler poiché ci costringe ad un approccio diverso da quello che potremmo avere (su altri siti) sfruttando le classi per individuare gli specifici elementi html che contengono i dati che ci interessano.

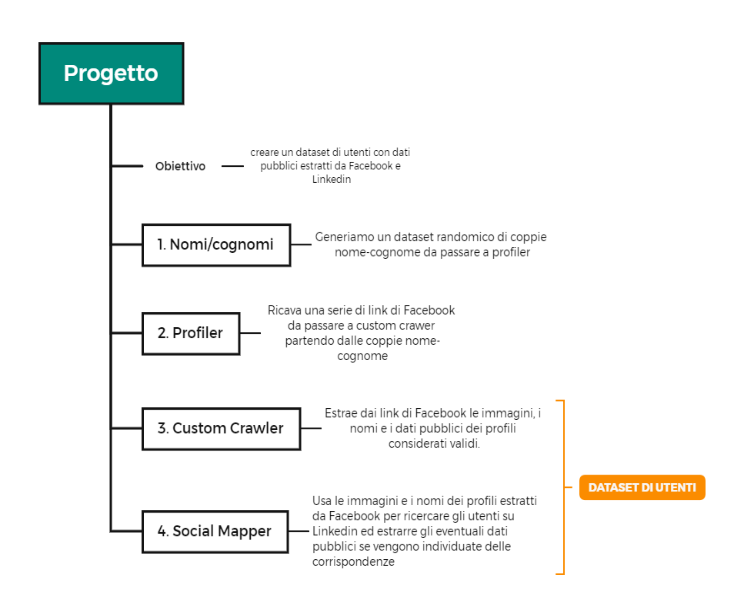


Figura 4: struttura del progetto

Dal diagramma in figura 4 è possibile osservare come è strutturato il progetto e la correlazione tra i vari strumenti utilizzati. Profiler permette di ricavare una serie di link a pagine Facebook prendendo in input un insieme di coppie nome-cognome che vengono generate randomicamente partendo dal dataset [19]. Il custom crawler sfrutta tali link per estrarre una lista di utenti di Facebook: per ogni utente il crawler estrae nome, cognome, immagine del profilo e i dati sensibili in chiaro sulla pagina. Immagine, nome e cognome di ogni profilo valido vengono poi passati a Social Mapper che si occupa di ricercare ed estrarre i dati sensibili di ogni utente su cui viene individuata una corrispondenza all’interno di Linkedin. I dati ottenuti dal custom crawler e da Social mapper verranno infine combinati su di un unico dataset.

Social Mapper non è uno strumento particolarmente preciso, e può individuare eventuali falsi positivi durante la ricerca. Sorge quindi la problematica di convalidare i dati ottenuti. Tale problematica è stata trattata solo in modo superficiale poiché buona parte dei falsi positivi di Linkedin possiedono un nome diverso dal profilo Facebook e, per tanto, è stato lineare escludere questi casi banali dal risultato finale. Per un controllo più accurato sarebbe necessario un confronto dei dati estratti tra i due social: due dati, ovvero lavoro e residenza, sono in chiaro su entrambi i profili e sono particolarmente adatti per tale scopo.

Sono stati analizzati un totale di 12039 link estratti a partire da 10000 coppie randomiche nome-cognome.

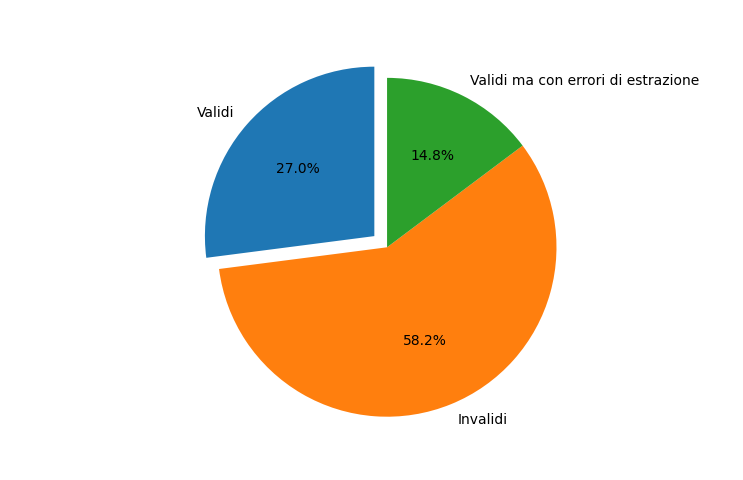


Figura 5: risultati di Profiler

Come mostrato dal grafico in figura 5, 3256 (27%) dei link corrispondono ad utenti validi mentre il restante risulta essere invalido o si è verificato un qualche errore nell’estrazione dell’immagine del profilo. Come utente valido viene considerata qualsiasi pagina Facebook avente un’immagine del profilo con uno e un solo individuo. Viene sfruttata la libreria dlib [17] per individuare il numero preciso di volti. Questo controllo è risultato necessario per poter, successivamente, passare le immagini a Social Mapper; senza tale restrizione, il numero degli utenti individuati sarebbe considerevolmente più alto.

# Analisi dei risultati

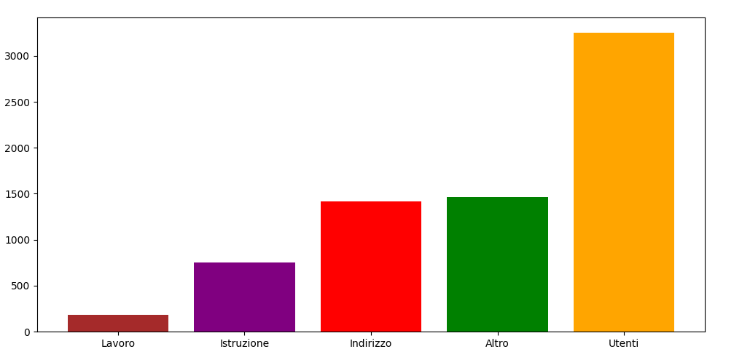
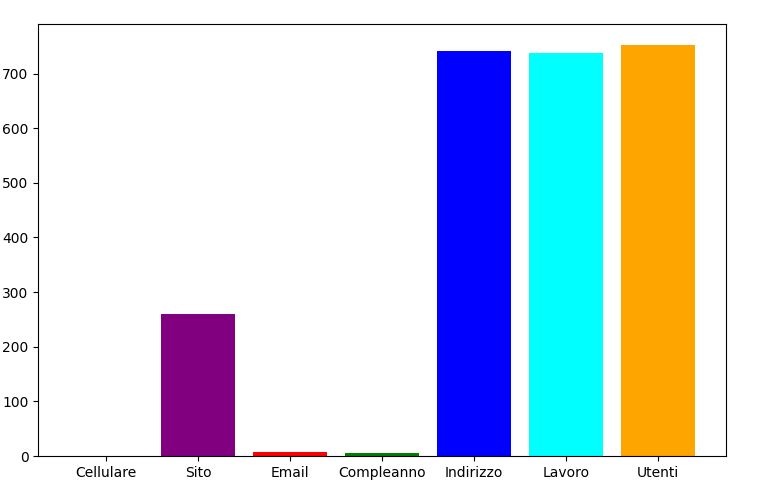
 Per quanto riguarda Facebook abbiamo tre informazioni in chiaro costanti su tutti i profili: lavoro, istruzione e indirizzo. Viene estratto anche un quarto dato in chiaro che sembra variare in ogni profilo: su alcuni troviamo il numero di followers, su altri la data di iscrizione al social.

Figura 6: numero di dati ottenuti da Facebook in relazione agli utenti totali

Come si osserva dal diagramma in figura 6, il dato più comune è il quarto dato variabile seguito dall’indirizzo che risulta essere il dato fisso inserito dal maggior numero di utenti. Non esiste un dato che la maggioranza degli utenti inserisce. Dal campione risulta che all’incirca il 48% (1567) degli utenti ha inserito almeno uno fra i quattro dati estratti, mentre da più della metà dei profili non viene estratto nessun dato in chiaro.

Da Linkedin potremmo potenzialmente estrarre un insieme di dati più ampio che include: cellulare, sito web personale, e-mail personale, compleanno, indirizzo e lavoro.

Figura 7: numero di dati ottenuti da Linkedin in relazione agli utenti totali

Come si osserva dal diagramma in figura 7, non viene estratto nessun numero di cellulare per tutti i 753 utenti per cui viene individuata una corrispondenza su Linkedin. Dal campione di utenti il numero di cellulare risulta quindi essere un dato privato di default e non direttamente estraibile tra le informazioni del profilo. Per quanto riguarda le altre informazioni, lavoro ed indirizzo risultano essere i dati più comuni che vengono inseriti dalla quasi totalità degli utenti. Il sito web personale viene inserito in alcuni casi e solo in piccole percentuali degli utenti troveremo e-mail e compleanno.

|  |  |  |  |
| --- | --- | --- | --- |
| Sito | Dato | Ricorrenze | Percentuale |
| Facebook | Lavoro | 185 | 5.6% |
| Facebook | Istruzione | 748 | 23% |
| Facebook | Indirizzo | 1420 | 43.6% |
| Facebook | Altro | 1465 | 45% |
| Linkedin | Cellulare | 0 | 0% |
| Linkedin | Sito | 260 | 34.5% |
| Linkedin | E-Mail | 8 | 1% |
| Linkedin | Compleanno | 6 | 0.8% |
| Linkedin | Indirizzo | 742 | 98.5% |
| Linkedin | Lavoro | 737 | 97.8% |

Figura 8: numero totale dei dati ottenuti e percentuale degli utenti che hanno tale dato in chiaro

Ricapitolando, il dataset finale sarà composto da un insieme di 3256 utenti. Per 753 di questi utenti è stato possibile individuare una corrispondenza su Linkedin. Il numero delle informazioni estratte è riportato dalla tabella in figura 8.

# Conclusione e sviluppi futuri

Dal lavoro svolto è possibile giungere ad una serie di conclusioni:

* La quantità di dati in chiaro ricavabili dai social networks Facebook e Linkedin
* Le percentuali di tali dati che è possibile ricavare mediamente da un campione randomico di utenti
* Che i social network sono in continua evoluzione e sono necessari degli strumenti migliori per la ricerca in tale campo.

Come risultati finali avremo, oltre al dataset, anche il codice utilizzato per l’estrazione dei dati sensibili. Tale codice non è però supportato e non sarà efficace sul lungo termine. Uno dei possibili sviluppi futuri sarebbe quindi quello di realizzare dei sistemi più efficaci sotto questo aspetto; il dataset creato può servire da campione di utenti per questo e per altri lavori di ricerca. Concludendo, un altro sviluppo particolarmente utile all’ambito sarebbe quello di affrontare, nello specifico, l’analisi cross-social e la convalida delle informazioni e dei profili analizzati: sarebbero utili metodologie precise sia per, ad esempio, distinguere utenti reali da possibili utenti fake, sia per verificare se due o più profili individuati su social diversi appartengano effettivamente allo stesso individuo.

# Riferimenti

1. Brian Dean, https://backlinko.com/social-media-users#social-media-usage-stats, 2021
2. Dragan Savić, Mladen Veinović, “**Challenges of general data protection regulation (GDPR)**”Singidunum University, Belgrade, Serbia, 2018
3. Sangeeta Kumari, Shailendra Singh, “**A Critical Analysis of Privacy and Security on Social Media**”, Computer Engineering & Application National Institute of Technical, Bhopal, India, 2015.
4. Jie Tang, Duo Zhang, and Limin Yao, “**Social Network Extraction of Academic Researchers**” Department of Computer Science and Technology, Tsinghua University
5. Jain, A.K., Sahoo, S.R. & Kaubiyal, J. “**Online social networks security and privacy: comprehensive review and analysis**”. Complex Intell. Syst. 7, 2157–2177 (2021), <https://doi.org/10.1007/s40747-021-00409-7>
6. Abawajy, Jemal H., Mohd Izuan Hafez Ninggal and Tutut Herawan. “**Privacy Preserving Social Network Data Publication**” IEEE Communications Surveys & Tutorials 18 (2016): 1974-1997.
7. Adrienne Felt, David Evans, “**Privacy Protection for Social Networking Platforms**”, University of Virginia, Oakland, CA. 22 May 2008.
8. David J. Houghton & Adam N. Joinson (2010) Privacy, “**Social Network Sites, and Social Relations**”, Journal of Technology in Human Services, 28:1-2, 74-94, DOI: 10.1080/15228831003770775
9. Sarikakis, Katharine, and Lisa Winter. “**Social Media Users’ Legal Consciousness About Privacy**.”, (January 2017). <https://doi.org/10.1177/2056305117695325>.
10. Lusiana Citra Dewia , Meilianaa, Alvin Chandraa , “**Social Media Web Scraping using Social Media Developers API and Regex**”, Procedia Computer Science, 2019, Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia, 11480
11. Batrinca, B., Treleaven, P.C., “**Social media analytics: a survey of techniques, tools and platforms**”, AI & Soc 30, 89–116 (2015). <https://doi.org/10.1007/s00146-014-0549-4>
12. Stefan Stieglitz, Milad Mirbabaie, Björn Ross, Christoph Neuberger, “**Social media analytics – Challenges in topic discovery, data collection, and data preparation**”, International Journal of Information Management, Volume 39, 2018
13. Di Minin, Enrico, Christoph Fink, Anna Hausmann, Jens Kremer, and Ritwik Kulkarni. “**How to Address Data Privacy Concerns When Using Social Media Data in Conservation Science**” Conservation Biology 35, no. 2 (March 22, 2021): 437–446. doi:10.1111/cobi.13708.
14. Erik Norvelle, “**NamesDatabases**”, https://github.com/smashew/NameDatabases/tree/master/NamesDatabases
15. D. E. King, “**Dlib-ml: A Machine Learning Toolkit**”, Journal of Machine Learning Research , 10, 1755–1758, 2009
16. Greenwolf, “**Social Mapper**”, https://github.com/Greenwolf/social\_mapper
17. Rog3rSm1th, “**Profil3**r”, <https://github.com/Rog3rSm1th/Profil3r>
18. “**Selenium**”, <https://www.selenium.dev/>
19. “**Dlib**”, https://github.com/davisking/dlib
20. “**BeautifulSoup**”, https://www.crummy.com/software/BeautifulSoup/bs4/doc/