



UNIVERSITA' POLITECNICA DELLE MARCHE
FACOLTA' DI INGEGNERIA

Corso di Laurea triennale in
Ingegneria Informatica e dell'Automazione

ACQUISIZIONE E ANALISI DEI POST DI FACEBOOK
FACEBOOK POSTS ACQUISITION AND ANALYSIS

Relatore:

Prof. Luca Spalazzi

Tesi di Laurea di:

Samuel Civitarese

Correlatore:

Prof. Francesco Pagliarecci

A.A. 2020/2021

Sommario

Sommario	- 3 -
Indice Figure.....	- 5 -
Introduzione.....	- 7 -
1. Fonte dei dati.....	- 9 -
1.1 Facebook.....	- 12 -
1.1.1 Struttura dei post	- 13 -
1.1.2 Commenti	- 17 -
1.2 Twitter.....	- 19 -
1.2.1 Twitter API.....	- 21 -
2. Facebook API	- 23 -
2.1 API Graph	- 23 -
2.2 Facebook Login	- 28 -
2.2.1 Autenticazione.....	- 29 -
2.2.2 Token di accesso	- 30 -
2.2.3 Autorizzazioni	- 32 -
2.2.4 Rate limit	- 35 -
3. Obiettivi software	- 37 -
3.1 Monitoraggio pagine.....	- 37 -
3.2 Discovery pagine	- 39 -
3.2.1 Esempio discovery pagine.....	- 40 -
4. Struttura del software	- 42 -
4.1 Fase di grabbing (Fase 1).....	- 43 -
4.1.1 Acquisizione dati.....	- 43 -

4.1.2 Fase di grabbing (Fase 2)	- 49 -
4.2 Fase di discovery	- 51 -
4.3 Filtro.....	- 54 -
4.3.1 Filtro Hashtag	- 55 -
4.3.2 Filtro post livello	- 56 -
5. Test	- 59 -
Fase 1 (Grabbing).....	- 59 -
Fase 2 (Discovery e filtraggio).....	- 63 -
Differenze del test con le ultime API	- 64 -
Conclusioni.....	- 67 -
Appendici	- 69 -
A1. Tabella versioni API Facebook	- 69 -
A2. Lista dei brand monitorati e relativi account Facebook	- 70 -
Sitografia	- 78 -

Indice Figure

Figura 1-1: Numero di utenti presenti su piattaforme social [Hootsuite & We Are Social (2021), “Digital 2021 Global Digital Overview,”]	- 9 -
Figura 1-2: Principali canali di ricerca dei brand utilizzati dagli utenti [Hootsuite & We Are Social (2021), “Digital 2021 Global Digital Overview”]	- 10 -
Figura 1-3: Classifica dei principali social network utilizzati [Hootsuite & We Are Social (2021), “Digital 2021 Global Digital Overview”]	- 11 -
Figura 1-4: Esempio di post su Facebook in cui si sono evidenziati gli elementi principali [https://www.facebook.com/nike/posts/10155790703238445]	- 13 -
Figura 1-5: Testo complessivo del post di Figura 1-4	- 14 -
Figura 1-6: Fotografia contenuta nel post di Figura 1-4	- 15 -
Figura 1-7: Numero di like e reazioni del post di Figura 1-4	- 15 -
Figura 1-8: Possibili reazioni che un utente può esprimere	- 16 -
Figura 1-9: Numero di commenti e condivisioni del post in Figura 1-4	- 16 -
Figura 1-10: Pulsanti di interazione	- 17 -
Figura 1-11: Esempio commenti relativi al post di Figura 1-4	- 18 -
Figura 1-12: Elenco dei possibili ordinamenti dei commenti	- 19 -
Figura 1-13: Esempio di tweet [https://twitter.com/armani/status/1405929518389501954?s=20].	- 20 -
Figura 2-1: Risultato della query.	- 25 -
Figura 2-2: Esempio dei campi “previous” e “next”.	- 27 -
Figura 2-3: Tool di esplorazione per la API Graph	- 30 -
Figura 2-4 nel rettangolo rosso sono evidenziate le autorizzazioni richieste dall'app.	- 31 -
Figura 3-1: Esempio di post su Facebook [https://www.facebook.com/nike/posts/10155790703238445]	- 38 -
Figura 4-1: Struttura del software [Andrea Lagone, “Ricerca ed analisi di tweet per il mondo della moda”]	- 42 -
Figura 4-2: Diagramma di flusso della procedura di acquisizione dati	- 44 -

Figura 4-3:Diagramma di flusso relativo alla routine di aggiornamento commenti	- 47
-	
Figura 4-4: Diagramma di flusso relativa alla fase di grabbing	- 49 -
Figura 4-5: Diagramma di flusso della fase di discovery (Fase 2)	- 53 -
Figura 4-6: Diagramma di flusso del Filtro Hashtag	- 56 -
Figura 4-7: Diagramma di flusso relativo al Filtro post livello	- 57 -
Figura 5-1: Risultati dei post acquisiti durante la Fase 1	- 60 -
Figura 5-2:Risultati dei brand acquisiti dopo la Fase 1	- 61 -
Figura 5-3:Risultati degli hashtag trovati nei commenti o nei post delle pagine di livello ≥ 1	- 62 -
Figura 5-4:Pagine taggate trovate dopo la Fase 1 (per privacy sono stati censurati i rispettivi id)	- 63 -
Figura 5-5: Lista delle pagine ammesse dopo la Fase 2	- 64 -
Figura 5-6: Pagina di prova utilizzata con la versione 11.0 dell'API Graph	- 65 -
Figura 5-7:Post acquisiti nella <i>Fase 1</i>	- 65 -
Figura 5-8:Commenti acquisiti nella <i>Fase 1</i>	- 65 -
Figura 5-9: Pagine taggate ricavate da post e commenti	- 65 -

Introduzione

Il lavoro esposto in questa tesi fa parte del progetto “**Social media sentiment analysis for fashion world**” proposto dalla ditta Fashion Bank di Ancona, a fine 2016, riguardante lo sviluppo di diversi software atti alla raccolta e al monitoraggio di informazioni presenti su riviste online, blog e social network riguardanti piccoli e grandi marchi, stilisti e tutte le possibili aziende operanti nel settore della moda. I dati acquisiti dovevano servire per la ricerca di nuovi dati d’interesse inerenti al mondo della moda e per analizzare le opinioni dei consumatori in merito ai prodotti di abbigliamento e alle tendenze del momento.

Il primo passo per la realizzazione dell’obiettivo principale dell’intero progetto, cioè quello della *Sentiment Analysis*, e quello di ottenere i dati d’interesse da una delle fonti sopracitate; per tale motivo mi è stato assegnato l’obiettivo di realizzare un software che fosse capace di ricavare le informazioni, da una lista prefissata di pagine, dal social network Facebook e, dall’analisi di tali informazioni ottenere potenziali pagine attinenti al mondo della moda.

I medesimi obiettivi sono stati assegnati al collega Andrea Langone che si è occupato della social network Twitter.

Di seguito è possibile trovare l’elenco dei contenuti trattati nei capitoli a seguire.

- Nel Capitolo 1 viene effettuata una breve introduzione sull’importanza dei social network e verranno analizzati le informazioni che è possibile ricavare da Facebook e Twitter.
- Il Capitolo 2 riguarderà il funzionamento dell’*API Graph* di Facebook, del meccanismo di autenticazione (*Facebook Login*) e autorizzazione per l’accesso alle funzionalità di suddette API.
- Nel Capitolo 3 si procederà all’introduzione degli obiettivi del software e del funzionamento delle fasi principali dell’applicazione.

- Nel Capitolo 4 viene descritta la struttura del software e di come vengono implementati gli algoritmi riguardanti le fasi di grabbing e di discovery.
- Nel Capitolo 5 si illustreranno i risultati ottenuti dall'applicazione rispettivamente con la versione 2.8 e 11.0 delle API di Facebook e le loro differenze.

Concludendo verranno illustrate le problematiche affrontate e gli obiettivi raggiunti dal lavoro esposto in questa tesi e di come, i recenti cambiamenti riguardo alle condizioni di utilizzo delle API, abbiano influito sugli obiettivi raggiunti. Inoltre, si suggeriranno eventuali migliorie che potranno essere apportate, in futuro, al software.

1. Fonte dei dati

Al giorno d'oggi il web presenta moltissime fonti di informazioni come testate giornalistiche, blog, biblioteche digitali, ecc. che mettono a disposizione dell'utente finale (persone fisiche, aziende, professionisti comunità, ecc.) un'ampia gamma di informazioni. Per accedere a tali fonti l'utente effettua delle ricerche utilizzando i motori di ricerca (Google, Yahoo, Bing, ecc.) che hanno lo scopo di indicizzare le pagine web, all'interno della rete, in modo da facilitare la ricerca di quest'ultime. Nel nostro caso è stato necessario trovare una tipologia di fonte di dati che fosse sempre aggiornata e che permettesse agli utenti di interagire con tali informazioni esprimendo le loro opinioni a riguardo; per tali motivazioni la scelta più naturale fu quella dei social network.

Tali piattaforme rappresentano vere e proprie comunità che permettono ai propri iscritti di ricercare informazioni, condividere contenuti (testuali o multimediali), acquistare o vendere oggetti, e numerose altre funzioni.

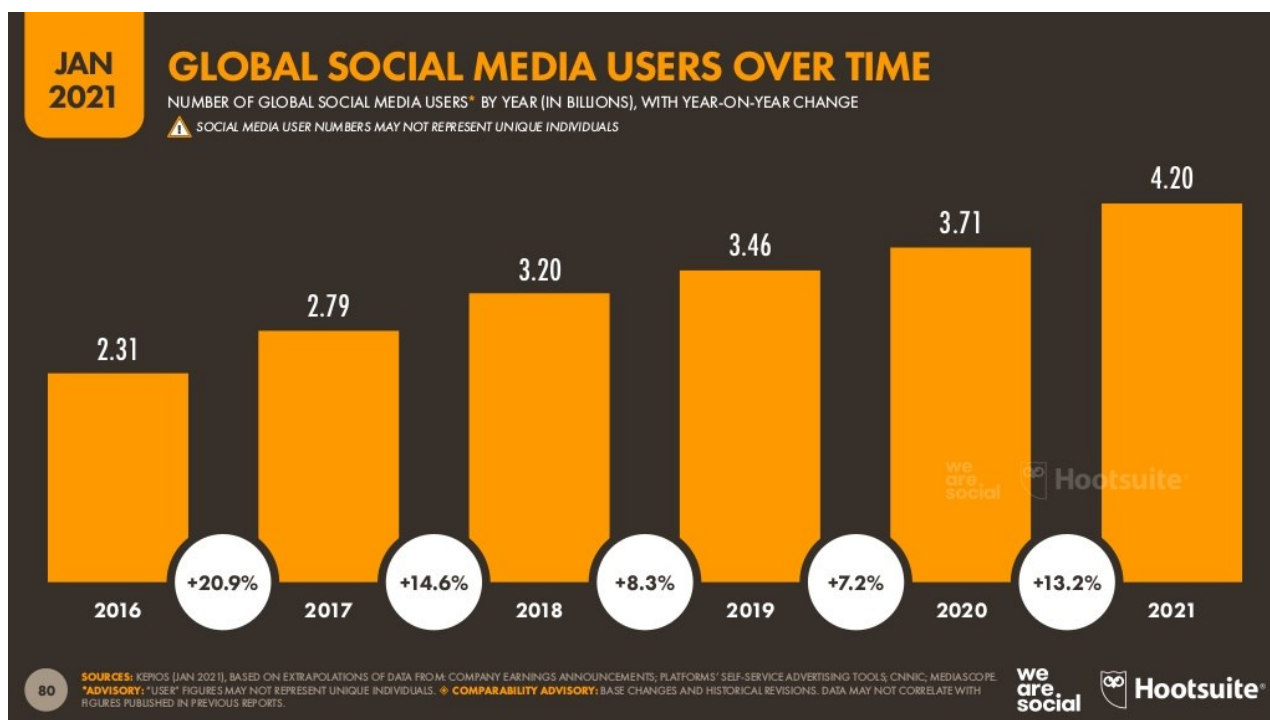


Figura 1-1: Numero di utenti presenti su piattaforme social [Hootsuite & We Are Social (2021), "Digital 2021 Global Digital Overview,"]

È interessante sottolineare, mettendo a confronto i dati dei report pubblicati nel 2016 [1] e nel 2021 [2] da *Hootsuite & We Are Social*, che il numero totale di persone che hanno accesso a internet è passato da 3,419 miliardi nel 2016 a circa di 4,66 miliardi nel 2021 (+23% rispetto al 2016) di cui 4,20 miliardi fanno uso di piattaforme social, che confrontati ai 2,31 miliardi nel 2016, fanno registrare un aumento del 82% in soli quattro anni.

Risulta quindi fondamentale per le aziende cercare di riuscire a raggiungere i consumatori attraverso i social network poiché tali piattaforme si stanno rivelando una valida alternativa anche ai classici motori di ricerca per la ricerca dei brand da parte degli utenti (Figura 1-2).

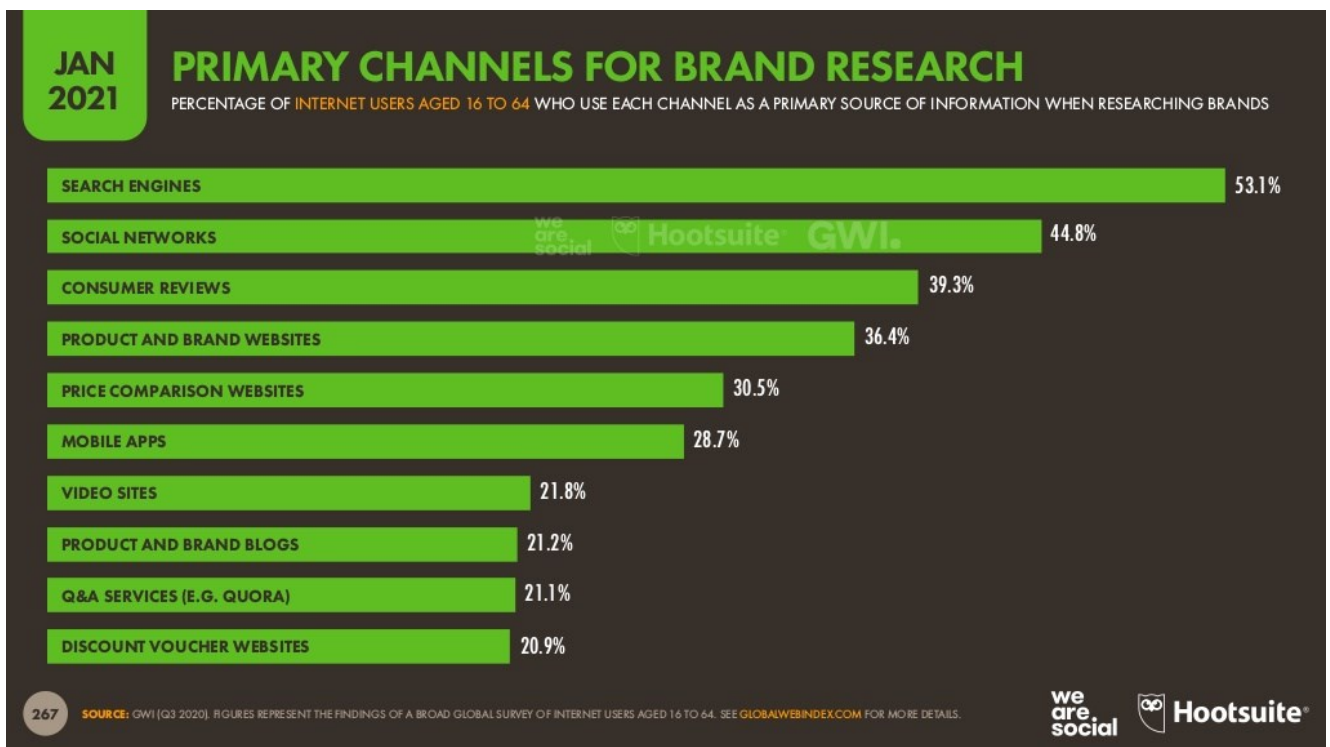


Figura 1-2: Principali canali di ricerca dei brand utilizzati dagli utenti [Hootsuite & We Are Social (2021), “Digital 2021 Global Digital Overview”]

Nell’ambito della *Sentiment Analysis* è di cruciale importanza poter disporre di un’ampia quantità di informazioni, allo stesso tempo aggiornate, al fine di poter studiare il sentimento dei consumatori.

I dati sopraesposti non fanno che rafforzare la scelta che nel 2016 ci ha portato a scegliere Facebook e Twitter come primarie fonti di informazioni. Considerando le ultime statistiche riguardanti le piattaforme social più utilizzate (Figura 1-3),

escludendo quelle relativi a servizi di messagistica (es: Whatsapp, Facebook Messenger, WeChat, Telegram, ecc.), Facebook e Twitter risultano essere ancora ampiamente utilizzati pur subendo una perdita di utenti dopo l'ingresso di nuove piattaforme, più user-friendly, come Instagram e TikTok.

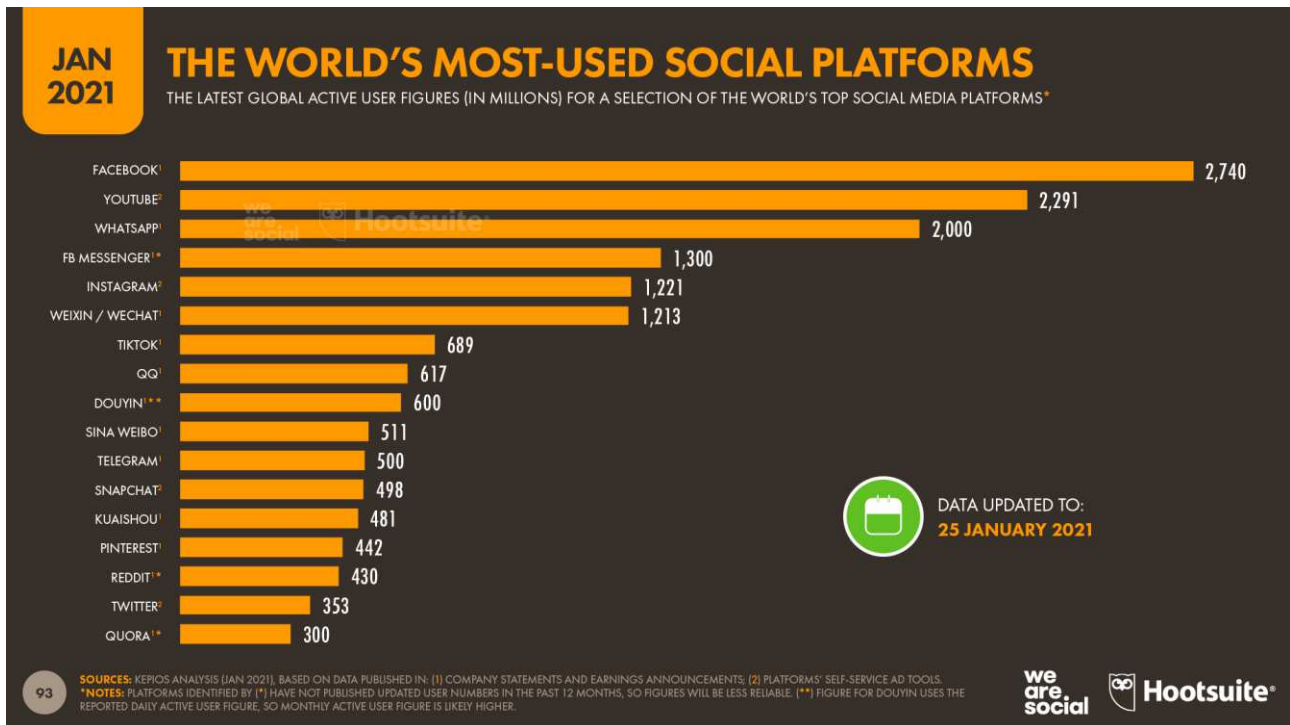


Figura 1-3: Classifica dei principali social network utilizzati [Hootsuite & We Are Social (2021), “Digital 2021 Global Digital Overview”]

Nei paragrafi successivi vado ad illustrare le caratteristiche salienti dei due social network in modo da descriverne il loro funzionamento e i tipi di dati che possono essere presi in considerazione nel nostro studio.

In particolare, lo studio relativo all’acquisizione dei dati da Twitter è stato svolto da Andrea Langone [3] e per tale motivo si procederà solamente a una breve descrizione della stessa, al fine di evidenziare le differenze tra le due piattaforme e di come vengono acquisiti i dati per le finalità di progetto.

1.1 Facebook

Facebook è un social network fondato nel 2004, di proprietà della società Facebook Inc., disponibile in più di 100 lingue su un'ampia gamma di dispositivi (smartphone, tablet, smart TV, ecc.), permette agli utenti la creazione di un profilo gratuito fornendo i propri dati personali come nome, cognome, data di nascita e indirizzo e-mail [4]. È possibile aggiungere ulteriori informazioni come luogo di nascita, città di nascita e residenza, posizione lavorativa, interessi personali, situazione sentimentale, ecc., che permettono alla piattaforma di suggerire all'utente potenziali amicizie, interessi e pubblicità mirate.

Le tipologie di servizi messi a disposizione da Facebook sono due: profilo personale e pagina pubblica.

Il profilo personale, chiamato anche diario, permette all'utente privato di pubblicare foto, video o annunci testuali mentre le pagine pubbliche vengono messe a disposizione di aziende, brand, organizzazioni e personaggi pubblici in modo che possano condividere notizie, eventi o qualsiasi altra informazione con la propria platea di utenti.

Ciò che permette di mettere in relazione un utente con una pagina, o con qualsiasi altro contenuto, è il tasto *mi piace (like)* che consente agli utenti di esprimere il loro apprezzamento o anche una reazione emotiva tramite sei emoticon¹ differenti.

Similarmente le relazioni tra utenti sono rappresentate tramite le *amicizie*, che permette all'utente di espandere la propria rete sociale fino ad un massimo di 5000 amici.

Ogni utente ha a disposizione sulla propria homepage un aggregatore di notizie, chiamato *news-feed*, in cui vengono mostrati i contenuti pubblicati da amici, pagine o gruppi (pubblici o privati) e inserzioni pubblicitarie.

¹ Le emoticon, faccine o smiley sono riproduzioni stilizzate delle principali espressioni facciali umane che esprimono un'emozione.

1.1.1 Struttura dei post




Figura 1-4: Esempio di post su Facebook in cui si sono evidenziati gli elementi principali

[\[https://www.facebook.com/nike/posts/10155790703238445\]](https://www.facebook.com/nike/posts/10155790703238445)

Qualsiasi contenuto che viene pubblicato (immagini, video, testi, link, evento, album, ecc.) su Facebook viene inserito in un *post* (Figura 1-4) e all'interno di esso possiamo trovare un certo numero di componenti basilari:

- 1) post: evidenziato con il rettangolo rosso tratteggiato, può essere di diversi tipi e allo stesso tempo può contenere anche elementi diversi come nel caso di figura

in cui si ha un testo iniziale seguito da un'immagine;

2) Autore: fornisce informazioni circa l'autore del post, la data di pubblicazione, la posizione da dove viene scritto e la sua visibilità (indica quali utenti sono autorizzati a vederlo). Vicino al nome dell'autore è possibile trovare un simbolo () indicante che l'account è stato verificato dalla piattaforma. Solitamente gli account verificati sono quelli che rappresentano figure pubbliche, aziende o brand famosi.

3) Testo:

Outplay yourself.

[Roger Federer](#), Melbourne's defending champion and holder of 19 major titles, just beat his own record, by winning 20. [#justdoit](#)

Figura 1-5: Testo complessivo del post di Figura 1-4

Rappresenta tutto ciò che viene scritto all'interno del post, esso può contenere diversi elementi ipertestuali come hashtag, tag (citazioni), collegamenti esterni o emoticon.

4) Tag: [Roger Federer](#) (Figura 1-5).

L'obiettivo di tale elemento è quello di collegare in modo diretto il contenuto del post con un utente o una determinata pagina tramite un collegamento ipertestuale che rimanda al profilo citato dal tag. Nel caso che un utente o una pagina sono menzionati all'interno di un post, essi riceveranno istantaneamente una notifica di avviso.

5) Hashtag: [#justdoit](#) (Figura 1-5).

L'hashtag è una tipologia di tag identificato per mezzo del simbolo cancelletto # davanti a una parola o una frase.

La sua funzione è quella di aggregare post di uno stesso tema o contenuto, creati da profili differenti, e mostrarli all'utente all'interno di un news-feed

dedicato.

6) Contenuti multimediali: fotografia di Roger Federer (Figura 1-6)



Figura 1-6: Fotografia contenuta nel post di Figura 1-4

Lo spazio delimitato dal rettangolo 6, tratteggiato in giallo e sottostante al testo del post, può contenere diverse tipologie di elementi multimediali come foto, video, album (insieme di fotografie e/o video), ecc..

Nel caso in cui il testo non fosse presente il contenuto multimediale va a occupare tutto lo spazio a sua disposizione andando a descrivere in maniera visiva il contenuto del post.

7) Like e reazioni:



Figura 1-7: Numero di like e reazioni del post di Figura 1-4

Indica il numero totale di like e reazioni (stato d'animo degli utenti) ricevuti da un post o commento. Ogni post può ricevere likes e, eventualmente, reazioni che vengono mostrati all'interno del riquadro 7 di Figura 1-4.

Nel nostro caso accanto al simbolo like viene mostrato un simbolo a forma di cuore (love) che sta ad indicare un tipo di reazione aggiuntiva che gli utenti hanno espresso riguardo al post. Nel nostro caso, il numero totale (48,368) è la somma dei like e delle reazioni love.

Facebook mette a disposizione sei tipi di reazioni emotive (escludendo il *mi piace*) che sono: Love, Ahah, Wow, Sigh o Grrr (Figura 1-8).

- Love: rispetto al classico *mi piace* rappresenta emozioni più intense come l'amore.
- Ahah: classica risata che ha un cenno di divertimento, allegria o comicità più che di apprezzamento.
- Wow: può assumere connotazioni positive, con gli aggettivi wow, fantastico, geniale, sensazionale, o connotazioni negative come dire "stai scherzando, vero?".
- Sigh: esprime tristezza o compassione verso il prossimo.
- Grrr: rappresentano uno stato d'animo furioso e viene utilizzato per esprimere disaccordo o disappunto. Viene spesso identificato dagli utenti come *non mi piace*.



Figura 1-8: Possibili reazioni che un utente può esprimere

8) Commenti e condivisioni:

Commenti: 9203 Condivisioni: 3272

Figura 1-9: Numero di commenti e condivisioni del post in Figura 1-4

Indica il numero totale di utenti e/o pagine che hanno commentato (campo commenti) il post o condiviso (campo condivisioni) sul proprio profilo o su profili altrui.

9) Pulsanti d'interazione:



Figura 1-10: Pulsanti di interazione

I pulsanti d'interazione (Figura 1-10) permettono all'utente di interagire con il post desiderato tramite tre azioni:

- *Mi Piace*: l'utente può esprimere la sua reazione tra le opzioni elencate nel punto 7.
- *Commenta*: utilizzato per commentare direttamente il post considerato.
- *Condividi*: permette agli utenti di condividere il post nel proprio diario o nel diario di un altro profilo purché in possesso delle autorizzazioni necessarie. In quest'ultimo caso l'utente che riceve il contenuto condiviso può accettare o meno la pubblicazione sul suo profilo.

1.1.2 Commenti

Abbiamo appena visto che è possibile commentare un post esprimendo le proprie opinioni o rispondere a commenti altrui.

Consideriamo la Figura 1-11 nella quale viene mostrato un estratto dei commenti relativi al post di Figura 1-4; l'intera struttura dei commenti è contenuta all'interno del rettangolo n.1, dove i commenti sono ordinati di default in base ai più pertinenti ma è possibile cambiare tale ordine facendo clic su *tutti i commenti* come mostrato in Figura 1-12.

Nel rettangolo n.2 è presente una casella di testo nella quale l'utente può inserire il proprio commento e, così come per i post, citare altri utenti e/o pagine, inserire hashtag o collegamenti a siti esterni; tramite i tre pulsanti laterali è possibile aggiungere contenuti extra come emoticon, gif o adesivi (stickers).

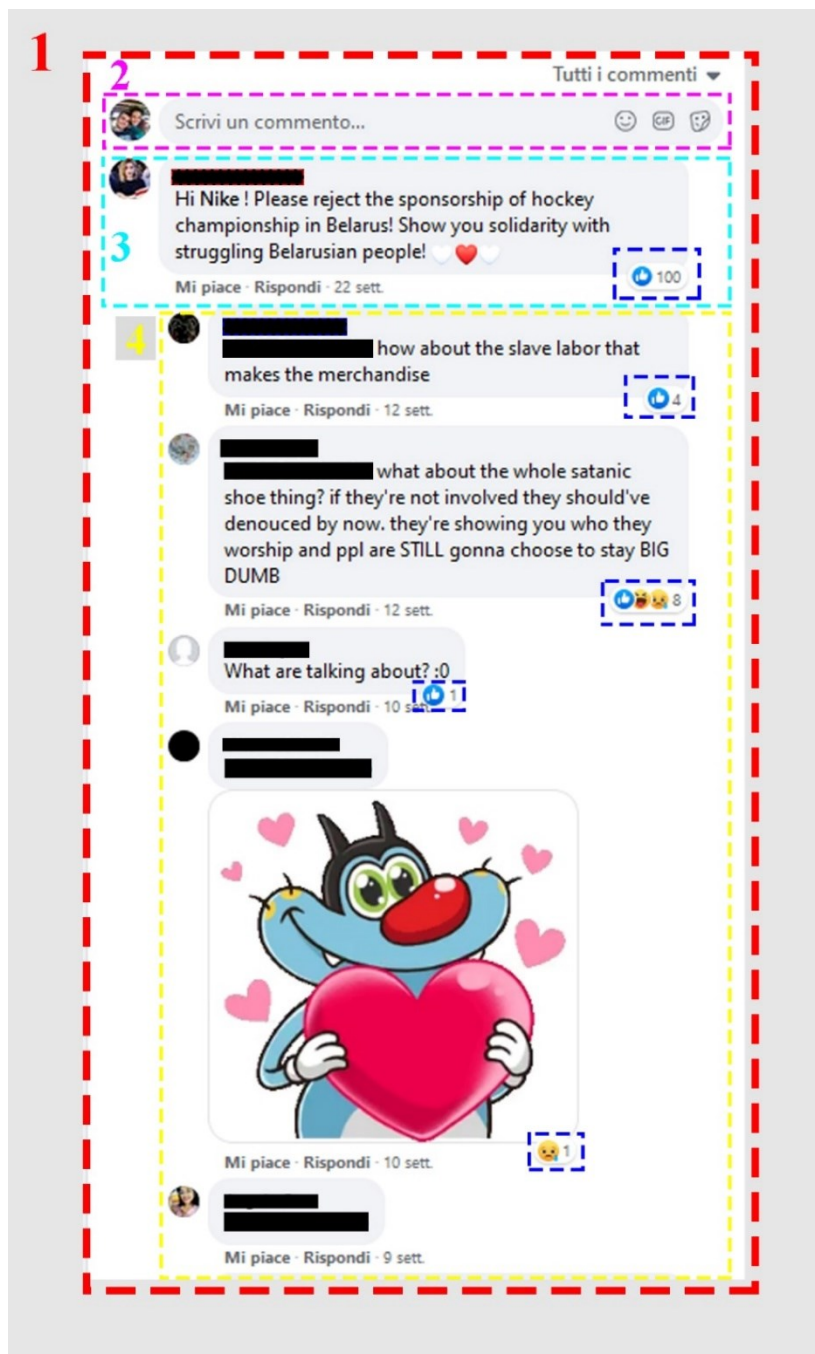


Figura 1-11: Esempio commenti relativi al post di Figura 1-4

Nello spazio evidenziato dal rettangolo n.3 possiamo notare un commento² primario con un tag all'autore (Nike in grassetto), del testo e delle emoticon finali.

² Per privacy i nomi degli utenti sono stati censurati per mezzo di rettangoli neri.
Anno Accademico 2020/2021

Al di sotto troviamo il rettangolo n.4, rientrato rispetto al precedente, in cui sono presenti i commenti, di altri utenti, in risposta a quello precedente.

Da notare che non tutti i commenti contengono del testo (come il penultimo) dato che Facebook consente di commentare anche con emoticon, gif animate o adesivi.

Infine, è possibile constatare, nei riquadri tratteggiati in blu, le reazioni ai singoli commenti.

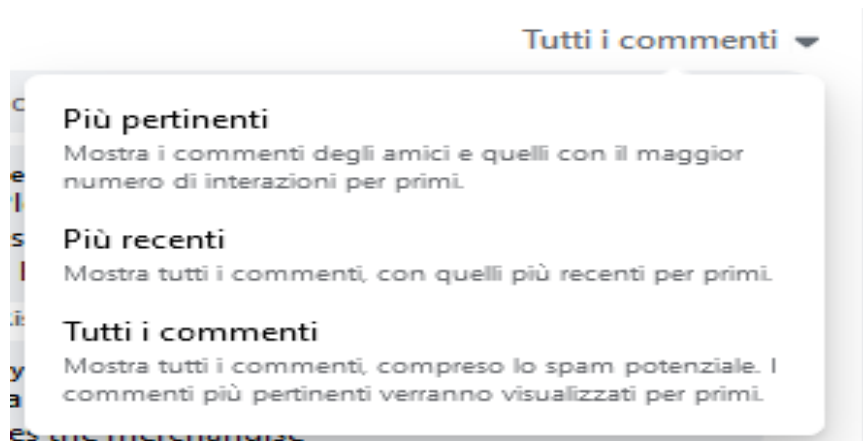


Figura 1-12: Elenco dei possibili ordinamenti dei commenti

1.2 Twitter

Twitter è un servizio di notizie e microblogging fondato nel 2006 e di proprietà della Twitter Inc. [5] e dopo Facebook è stato uno dei primi social network distribuito su scala mondiale e anche uno dei più utilizzati.

La piattaforma è completamente gratuita e consente agli utenti di condividere dei messaggi, chiamati *tweet*³, di lunghezza massima pari a 280 caratteri (in origine il limite era di 140 caratteri).

Ogni utente riceve gli aggiornamenti dei messaggi pubblicati dagli altri utenti e quando ciò accade si dice che l'utente *segue* o diventa *follower* di un altro utente; l'insieme degli utenti seguiti si chiama *following*, viceversa l'insieme degli utenti da cui si è seguiti è chiamato *followers*.

³ Il termine "tweet" sta per "cinguettare" da cui poi deriva il nome del social network "Twitter".

Tra gli utenti non vige l'obbligo di reciprocità: un utente può seguire un altro utente senza che quest'ultimo faccia lo stesso e viceversa.

Il profilo di un utente può essere pubblico o protetto, in quest'ultimo caso i contenuti sono privati e ogni potenziale follower deve essere approvato dal proprietario del profilo prima di poter accedere ai suoi contenuti.

Twitter mette a disposizione, per ogni utente, un aggregatore di tweet chiamato *timeline* (o *cronologia principale*) dove viene mostrato il flusso di tweet pubblicati dagli utenti seguiti e possibili suggerimenti, di altri account da seguire, in base ai propri interessi.



Figura 1-13: Esempio di tweet [<https://twitter.com/armani/status/1405929518389501954?s=20>].

Similmente come accade per i post di Facebook anche i tweet possono contenere, oltre al testo, tag ad altri utenti, hashtag, immagini e/o video (Figura 1-13); inoltre l'utente può effettuare quattro operazioni:

- 1) esprimere il gradimento per il tweet mediante il pulsante a forma di cuore; rappresenta l'equivalente del *mi piace* di Facebook;
- 2) rispondere ad un tweet: in questo caso verrà pubblicato sul profilo di chi commenta il tweet originale e sotto di esso il relativo commento (non deve necessariamente essere in forma testuale); l'autore del tweet riceverà una notifica avvertendolo del commento.
- 3) Retwittare il tweet: permette all'utente di pubblicare sul proprio profilo il tweet originale (tranne se non risulta essere stato protetto dall'autore) in modo da mostrarlo ad ogni suo follower, permettendo loro di interagire a sua volta con esso. A differenza della risposta non viene aggiunto nessun commento.
- 4) Condividi: tramite questo pulsante si ha accesso a tre funzioni aggiuntive che permettono all'utente di inviare il tweet tramite messaggio privato, aggiungerlo ai preferiti o copiare il suo URL.

Per il tweet di esempio (Figura 1-13) abbiamo rispettivamente 3 risposte, 30 retweet e 186 like.

I commenti e i retweet godono, di default, di visibilità pubblica (tranne per gli account protetti), e ogni utente può vedere quali sono gli utenti che hanno commentato o retwittato un determinato tweet senza essere loro follower.

1.2.1 Twitter API

Twitter fornisce diverse tipologie di API [6] che permettono di interagire con la piattaforma nelle medesime modalità previste per l'utente finale.

Gli endpoint sono raggruppabili in cinque categorie principali [7]:

1. account e utenti;
2. tweet e risposte;
3. messaggi Diretti;

4. annunci;
5. strumenti e SDK per i publisher.

Twitter mette a disposizione tre tipologie di API:

1. standard: permettono di accedere, gratuitamente, a dati e operazioni che normalmente un utente può eseguire sulla piattaforma.
Si hanno le seguenti limitazioni: dati disponibili (es. non sono presenti tutti i tweet della piattaforma), un solo tweet restituito per chiamata, rate limit [8] basso (numero di richieste in un data finestra temporale) e query con opzioni limitate.
2. Premium: a pagamento, si basano su una versione limitata delle API enterprise che forniscono una serie di vantaggi come l'accesso ai dati degli ultimi 30 giorni o, se l'endpoint lo permette, l'intera cronologia, rate limit più elevato rispetto la versione standard, query più complesse, più tweet restituiti per singola richiesta, metadati aggiuntivi e informazioni di geolocalizzazione avanzate.
3. Enterprise: versione completa, a pagamento, con minori limitazioni e supporto con console dedicate di sviluppo e monitoraggio.

Il software realizzato da Andre Langone utilizza la versione standard delle API e i dati che vengono acquisiti sono: testo del tweet, URL di immagini e/o video, URL esterni, geolocalizzazione (latitudine e longitudine) del tweet, hashtag e utenti taggati. Tali dati verranno integrati a quelli ottenuti da Facebook per perseguire gli obiettivi indicati nel capitolo 3. .

2. Facebook API

2.1 API Graph

L'*API Graph* [9] è lo strumento principale che Facebook offre per l'integrazione delle applicazioni con le informazioni contenute all'interno del social network. Tali API permettono sia di ottenere informazioni presenti sulla piattaforma e sia di interagire con essa utilizzando tutte le funzionalità che normalmente vengono messe a disposizione dell'utente (per esempio: pubblicare un post o delle fotografie, rispondere a dei commenti, ecc.).

Come nella maggior parte dei servizi Web odierni tali API seguono i principi dell'architettura REST (REpresentational State Transfer) basata sul protocollo HTTP [10].

Tale approccio architetturale prevede che il client richieda una specifica risorsa ad un determinato indirizzo URL (tramite i metodi GET e POST del protocollo HTTP) e il server, dopo aver elaborato la richiesta, trasferisce al richiedente lo stato che rappresenta la risorsa in un formato predefinito che nel caso dell'API Graph è di tipo JSON.

L'API Graph si basa sull'idea di *social graph* ovvero quella di rappresentare Facebook come un grafo costituito da tre elementi fondamentali:

- nodi (nodes): rappresentano i singoli oggetti come ad esempio utenti, pagine, post e commenti;
- segmenti (edges): sono i collegamenti (archi) fra raccolte di nodi (fotografie contenute in un album) o singoli nodi (like di un utente ad una pagina);
- campi (fields): dati contenuti in un oggetto come il nome di una pagina o il testo di un commento.

Quasi tutte le chiamate, di lettura o scrittura, vengono inoltrate all'host *graph.facebook.com* tranne che per l'upload di video dove l'host di riferimento è *graph-video.facebook.com*.

Ciò significa che l'indirizzo base a cui effettuare le chiamate sarà del tipo *https://graph.facebook.com/* o *http://graph-video.facebook.com* a seconda del caso di utilizzo.

Prima di poter effettuare un qualsiasi tipo di richiesta (lettura o scrittura) bisogna in qualche modo conoscere la posizione dell'oggetto all'interno del *social graph*: non basta infatti conoscere il nome dell'oggetto (es. il nome di una pagina o di un utente) poiché potrebbero essere presenti altri oggetti con lo stesso nome.

Per tale motivo ogni nodo, o oggetto, viene identificato all'interno del social network da un codice identificativo univoco che all'atto della chiamata verrà specificato nel campo "id-nodo" all'interno dell'URL *https://graph.facebook.com/id-nodo* .

Eseguendo una chiamata di questo tipo, senza ulteriori specifiche, verrà restituito un numero predefinito di proprietà (campi) di quell'oggetto.

Ad esempio, effettuando una query ad una generica pagina:

```
https://graph.facebook.com/id_pagina?access_token={access_token}
{
  "name": "Nome Pagina",
  "id": "id della pagina"
}
```

otteniamo come risultato i campi di default "name" e "id" (entrambi di tipo String) che indicano rispettivamente il nome della pagina e il suo id.

Nel caso in cui volessimo mostrare solamente le proprietà che ci interessano, dobbiamo specificare con l'ausilio del termine *fields* i nomi dei campi desiderati separati da virgole; i campi restituiti possono essere tipi primitivi (int, float, bool, ecc.) o oggetti.

Oltre alle proprietà di un oggetto è possibile specificare nel campo *fields* i segmenti ad esso associato. I segmenti restituiscono collezioni di nodi ad essi collegati che presentano a loro volta dei campi di default.

Analizziamo la seguente query:

https://graph.facebook.com/102469758344933?fields=id,category,business,photos,feed,access_token={access_token}.

Nel campo fields immettiamo:

- id: restituisce l'id univoco della pagina (campo);
- category: categoria alla quale appartiene la pagina (campo);
- business: tipologia di business associata alla pagina (campo);
- photos: restituisce tutte le immagini pubblicate dalla pagina (segmento);
- feed: lista dei post e link pubblicati dalla pagina o dai visitatori sulla pagina (segmento).

```
{
  "id": "102469758344933",
  "category": "Negozio di abbigliamento maschile",
  "photos": {
    "data": [
      {
        "created_time": "2020-11-06T16:55:25+0000",
        "id": "102469791678263"
      }
    ],
    "paging": {
      "cursors": {
        "before": "QVFIU1BndmFqeWxoazBOcE9oSUd2pKkb2pVJB",
        "after": "QVFIU1BndmFqeWxoazBOcE9oSUdfUJkb2pZB"
      }
    }
  },
  "feed": {
    "data": [
      {
        "created_time": "2020-11-11T08:49:34+0000",
        "message": "Emporio Armani the best!",
        "id": "102469758344933_105784618013447"
      },
      {
        "created_time": "2020-11-06T16:57:18+0000",
        "message": "Post 2 - senza commenti",
        "id": "102469758344933_102470681678174"
      }
    ],
    "paging": {
      "cursors": {
        "before": "QVFIUldTX25tMVkxLWcwRFNnZAGx44amp",
        "after": "QVFIUk5XQlJBODhiVmdYU2tzbEUY2V1Vo"
      }
    }
  }
}
```

Figura 2-1: Risultato della query.

Nella Figura 2-1 è possibile osservare il risultato ottenuto dalla query individuando tre blocchi principali:

- I. rappresenta il risultato dei campi del nodo principale (id e category);
- II. risultato relativo al segmento “photos”;
- III. risultato relativo al segmento “feed”.

La prima cosa che è possibile notare è che non risulta essere presente il risultato relativo al campo “business” poiché la pagina non ha specificato nessuna informazione a riguardo. In tali circostanze l’API Graph omette l’informazione dal risultato globale.

I blocchi II e III presentano un array “data” che contiene una lista di oggetti specifici: nel segmento “photos” gli elementi dell’array sono oggetti di tipo Photo mentre nel segmento “feed” sono di tipo Post e Link.

In tali blocchi è possibile notare un altro array (blocco viola) chiamato “paging”: quando viene eseguita una chiamata a un nodo o a un segmento potrebbe capitare che il numero di risultati sia troppo elevato per poterlo restituire in un’unica volta, per tale motivo ogni risposta viene paginata di default.

L’impostazione predefinita prevede la paginazione basata su cursore che viene rappresentato da una stringa casuale di caratteri. Il cursore identifica un elemento specifico nella lista dei risultati e può essere di due tipi: “before” e “after”. Il cursore “before” punta all’inizio della pagina di dati restituita mentre “after” alla fine della pagina.

Per poter scorrere le pagine in avanti (indietro) basta porre il cursore “after” (“before”), che in tal caso è chiamato parametro, all’interno di una nuova query che conterrà sempre l’id del nodo da interrogare seguito dal campo o segmento interessato dalla paginazione:

[https://graph.facebook.com/102469758344933?fields=id,category,business,feed.fields\(after=’valor e cursore after’\)](https://graph.facebook.com/102469758344933?fields=id,category,business,feed.fields(after=’valor e cursore after’)) .

Come possiamo notare la query è identica alla precedente ma abbiamo aggiunto il termine `.fields(after="stringa casuale")` al segmento `feed`.

In realtà nel nostro esempio non c'è bisogno di effettuare una nuova query poiché l'array `"data"` contiene già tutti i risultati.

Per capire se i risultati sono contenuti tutti su una pagina, bisogna verificare se l'oggetto `paging` contiene al suo interno almeno uno di questi due campi facoltativi: `"next"` e `"previous"` (Figura 2-2); entrambi i campi contengono il link che restituisce la pagina di dati successiva, o precedente, relativo al nodo o al segmento considerato escludendo tutti gli altri contenuti nella query originaria (nel nostro caso non compariranno più i risultati relativi a `id`, `category` o `business` ma solo quelli relativi a `feed`).

```
{
  "data": [
    ...
  ],
  "paging": {
    "cursors": {
      "after": "stringa_after",
      "before": "stringa_before"
    },
    "previous": "https://graph.facebook.com/{id_nodo_o_segmento}/feed?before=stringa_before"
    "next": "https://graph.facebook.com/{id_nodo_o_segmento}/feed?after=stringa_after"
  }
}
```

Figura 2-2: Esempio dei campi `"previous"` e `"next"`.

L'utilizzo di questa funzionalità risulta essere conveniente poiché non è necessario di volta in volta memorizzare i cursori che risulterebbero obsoleti nelle chiamate successive.

Analogamente a quanto detto sulla paginazione basata su cursore, sono disponibili altre due tipologie di paginazione: una basata sul tempo e l'altra basata sull'offset.

La paginazione su base temporale permette di navigare tra i dati dei risultati utilizzando timestamps Unix che puntano ad istanti di tempo specifici nella lista dei risultati. In tal caso i parametri temporali caratteristici sono `"since"` e `"until"`, ed essi

puntano rispettivamente all'inizio o alla fine di una serie di dati.

Questa tipologia di paginazione permette dunque di ottenere una serie di dati compresa in un intervallo di tempo ben definito (differenza massima di sei mesi).

L'ultima tipologia è quella basata sull'offset che viene utilizzata quando si dà priorità al numero di oggetti restituiti e non all'ordine cronologico. In tal caso bisogna specificare nel parametro "offset" il numero di elementi da restituire a partire dall'inizio della pagina.

Bisogna tenere conto che il contenuto delle pagine potrebbe variare se vengono aggiunti o rimossi elementi alla lista che si sta paginando, ad esempio l'aggiunta in tempo reale di un post ad una pagina.

L'aspetto più importante da tenere in considerazione è che tale tipo di paginazione non è supportato per tutte le chiamate API e per tale motivo è consigliato l'uso di quella su cursore o su tempo.

Un ulteriore parametro comune ai tre tipi di paginazione è "limit" che permette di limitare il numero di elementi che potrebbero essere restituiti dall'interrogazione.

Parte essenziale di ogni chiamata effettuata all'API Graph è il parametro "access_token" in cui viene specificata una stringa alfanumerica che serve per identificare un utente, un'applicazione o una pagina e può essere usata dall'applicazione per le nostre chiamate.

2.2 Facebook Login

Facebook Login [11] è il sistema che il social network mette a disposizione per effettuare l'accesso sulla piattaforma attraverso diverse piattaforme come smartphone, computer, smart TV ed in tutti quei dispositivi dove sia possibile installare applicazioni o semplicemente utilizzando un browser web.

Tale sistema prevede due possibili scenari di utilizzo: l'autenticazione e la richiesta delle autorizzazioni per l'accesso ai dati degli utenti.

Nella maggior parte dei casi questi due scenari vengono presentati all'utente in maniera visiva dando la possibilità all'utente di inserire le proprie credenziali per l'accesso ed

in un secondo momento di dare l'autorizzazione ad una applicazione (per esempio un videogioco), presente sul social network, di accedere ai suoi dati personali.

2.2.1 Autenticazione

Facebook utilizza OpenID come sistema di autenticazione basato a sua volta sul protocollo Authorization 2.0 (OAuth 2.0).

Tale protocollo è descritto da tali fasi principali:

- il *client* (l'applicazione di terza parte) chiede all'utente (*Resource Owner*) di accedere al server di autorizzazione. Dopo l'accesso l'utente decide se accettare o rifiutare la richiesta del *client*.
- Se l'utente accetta tale richiesta del *client* esso viene reindirizzato al *client* con un codice di autorizzazione.
- Il *client* richiede un *token* di accesso al server di autorizzazione fornendo il codice ottenuto nella fase precedente.
- Il server di autorizzazione autentica il *client* e controlla il codice di autorizzazione e se risultasse valido risponde al *client* fornendo il *token* di accesso.
- Una volta ottenuto l'accesso *token* il *client* può richiedere i dati desiderati al *resource server*.

Per implementare Facebook Login all'interno della propria applicazione è sufficiente utilizzare appositi software development kit (SDK) sviluppati da Facebook ma nel nostro caso non risulta possibile poiché l'applicazione non è provvista di interfaccia grafica e le informazioni di accesso sono integrate direttamente all'interno dell'applicativo stesso. In questo caso l'accesso *token* viene generato dall'utente attraverso il tool di esplorazione per la API Graph e successivamente salvato all'interno del file di configurazione dell'applicazione.

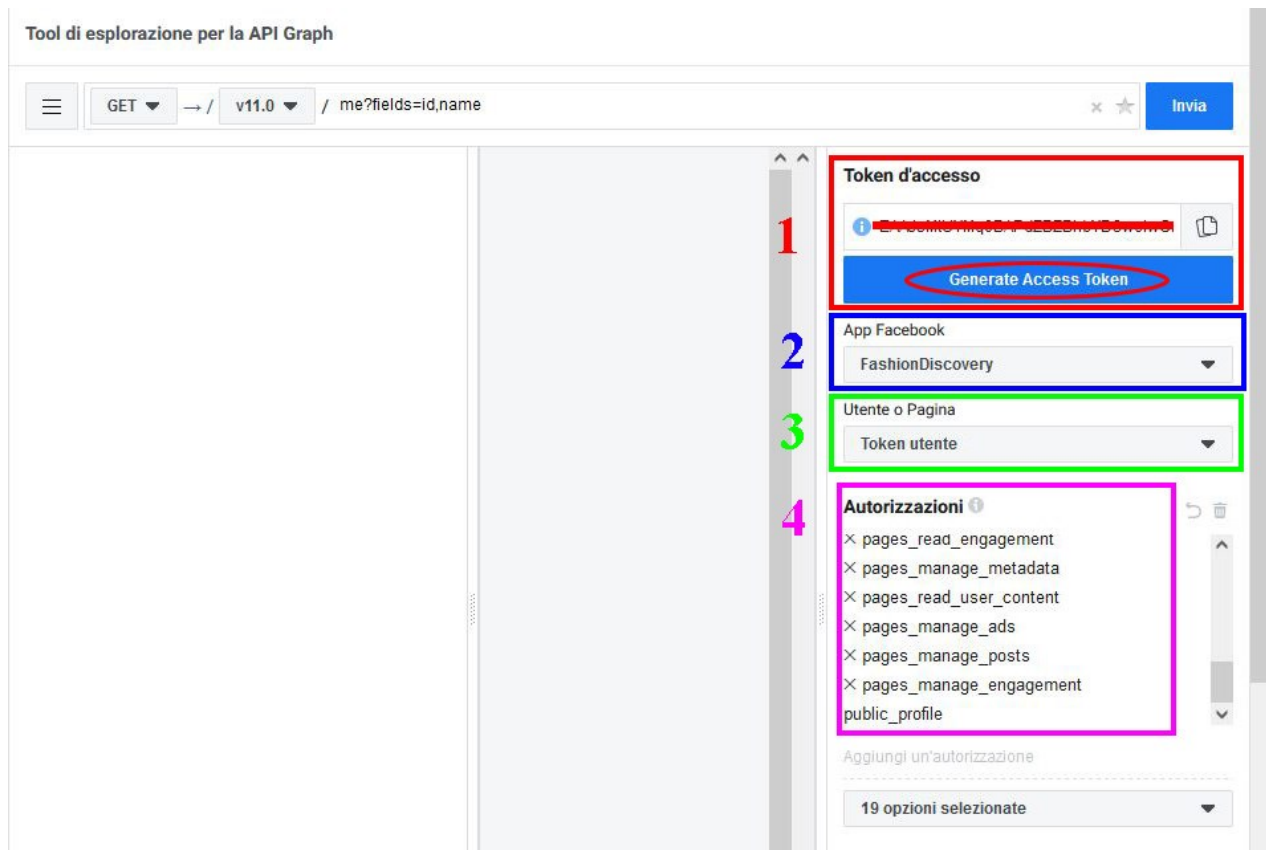


Figura 2-3: Tool di esplorazione per la API Graph

In Figura 2-3 è mostrato il Tool di esplorazione per la Api Graph che viene divisa in quattro blocchi principali:

- 1) Nella parte superiore è indicato il token di accesso attuale (se presente) per l'app indicata nel campo 2, inoltre tramite il pulsante *Generate Access Token* è possibile generare un nuovo access token;
- 2) in tale campo possiamo scegliere l'applicazione a cui ci stiamo riferendo;
- 3) indica la tipologia d'access token;
- 4) lista delle autorizzazioni selezionate per l'applicazione di riferimento.

2.2.2 Token di accesso

I token d'accesso possono essere di quattro tipi [12]:

- 1) token d'accesso utente (user access token): necessario per consentire la lettura, modifica o scrittura di dati Facebook di un utente da parte di un'applicazione (Figura 2-4).

- 2) token d'accesso dell'app (app access token): necessario per modificare e leggere le impostazioni dell'app, viene generato tramite una chiave segreta preconfigurata tra l'app e Facebook;
- 3) token d'accesso della pagina (page access token): simile alla tipologia per l'utente, permette all'applicazione di leggere e modificare i dati di una pagina Facebook. Per ottenere tale token bisogna essere in possesso di un token utente ed in seguito chiedere le autorizzazioni per le pagine;
- 4) token client: viene incorporato nei binari nativi dei client mobili o nelle app per computer. Esso è scarsamente utilizzato poiché ha un numero limitato di API alle quali può accedere e, ove possibile, viene consigliato l'uso degli altri tipi di token.

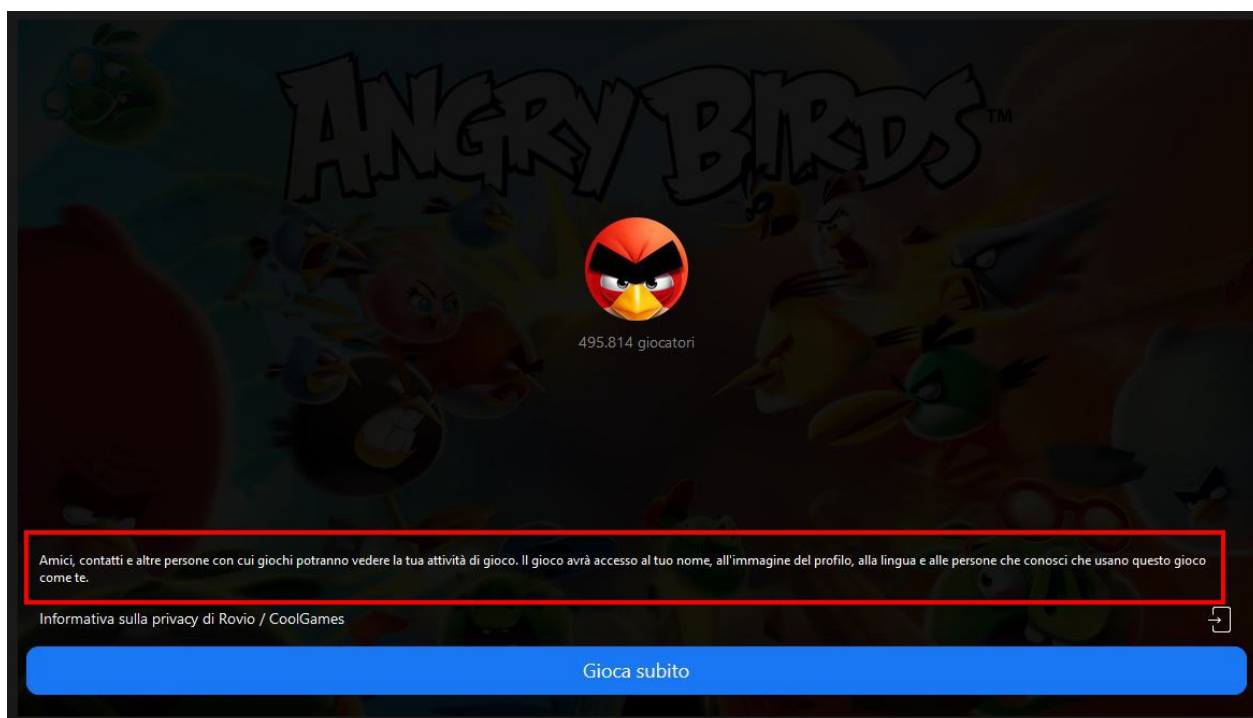


Figura 2-4 nel rettangolo rosso sono evidenziate le autorizzazioni richieste dall'app.

Le varie tipologie di token hanno una validità temporale limitata: essi possono durare un paio d'ore (token di breve durata) o circa sessanta giorni (token di lunga durata) ma tali valori sono indicativi e i token possono scadere in anticipo senza preavviso.

L'access token utente utilizzato dall'applicazione, generato manualmente tramite il tool di esplorazione per la API Graph, è un token di lunga durata.

Per evitare l'insorgere di errori ad ogni avvio dell'applicazione viene controllata la scadenza del token e se questa risulta essere inferiore a due giorni viene effettuata una richiesta di estensione della durata del token e, in caso di esito positivo, quest'ultimo viene sostituito con nuovo token fornito come risposta della chiamata precedente. Nel caso in cui ciò non dovesse essere possibile bisognerà generare un nuovo token in maniera manuale e salvarlo all'interno del file di configurazione del programma.

2.2.3 Autorizzazioni

Ogni tipologia di token d'accesso presenta un certo numero di endpoint dell'*API Graph* alla quale può accedere e uno o più endpoint possono fare parte di una specifica autorizzazione.

Le autorizzazioni essenziali fornite a tutte le applicazioni sono due:

1. *public_profile*: consente all'app di leggere i campi predefiniti del profilo pubblico su un nodo Utente;
2. *e-mail*: permette di leggere l'indirizzo e-mail principale di un utente.

Ogni altra tipologia di autorizzazione deve essere richiesta e successivamente approvata da Facebook. Sinteticamente possiamo riassumere in tre passi principali le operazioni da svolgere [13]:

1. vengono scelte le autorizzazioni di cui l'app ha bisogno escludendo quelle non necessarie;
2. l'app deve essere inviata per *l'Analisi delle app* [14] in modo che Facebook possa valutare che i dati forniti vengono utilizzati nei modi previsti e che la privacy degli utenti sia tutelata. Il livello di accesso

dell'autorizzazione passa da standard ad avanzato una volta ottenuti i permessi;

3. gli utenti che accedono all'app devono ricevere una richiesta per le autorizzazioni di cui l'app ha bisogno. Gli utenti devono poter approvare o negare tali autorizzazioni.

Per un'applicazione che viene eseguita su un server il secondo step non attuabile poiché risulterebbe difficile inviare l'app per essere analizzata. In questo caso Facebook mette a disposizione un'apposita procedura [15] che consiste in primo luogo di descrivere in dettaglio il funzionamento della propria applicazione, in secondo luogo di come i dati, forniti dal social network, vengono utilizzati e salvati. In ultimo luogo è necessario fornire un video di esempio nella quale viene mostrato il funzionamento dell'applicazione.

Le autorizzazioni di cui il nostro software necessita sono quelle relative alla lettura dei post pubblicati da una pagina ed i relativi commenti, sia che essi siano fatti dagli utenti o dalla pagina stessa.

Tali autorizzazioni sono contenute nella funzione **Page Public Content Access** [16] che consente all'applicazione di accedere all'*API Pages Search* (permettono la ricerca dei nodi all'interno del social network) e di leggere dati pubblici delle pagine per le quali non si dispone delle autorizzazioni **pages_read_engagement** e **pages_read_user_content** (disponibili, previa conferma, solo ai proprietari delle rispettive pagine).

L'accesso a tale funzione, e quindi ai contenuti pubblici delle pagine, non richiedeva particolari autorizzazioni con versioni dell'*API Graph* precedenti alla 3.0, introdotta solamente nel maggio 2018 [17] insieme alle procedure di App Review come risposta ai problemi di privacy sollevati con lo scandalo Facebook-Cambridge Analytica dove furono raccolti i dati personali di circa 87 milioni di persone senza il loro esplicito consenso.

Il programma descritto in questa tesi venne realizzato a dicembre 2016, utilizzando la versione 2.8 dell'*API Graph*, ben prima dell'introduzione del sistema di review delle

app, e la differenza con l'attuale versione (11.0) risiede principalmente come vengono ottenute le autorizzazioni. Se in precedenza era possibile ottenere libero accesso agli endpoint delle pagine ora è necessario che l'applicazione venga analizzata in modo da garantire la privacy degli utenti. Inoltre, l'uso della funzione **Page Public Content Access** è ristretto solamente agli account di tipo aziendale (business) verificati attraverso il processo di *Business Verification* [18] che consiste nell'associare l'identità dell'account ad un'entità aziendale realmente esistente. I dati forniti da tale funzione devono essere privi di riferimenti personali o anonimizzate (non deve essere possibile ricavare informazioni personali da tali dati) e gli intenti di utilizzo sono ristretti all'analisi, benchmarking competitivi, finalità di marketing o pubblicitarie. Nel caso in cui non si disponesse di un account aziendale non è possibile accedere ai dati pubblici delle pagine che non sono di proprietà dell'utente che utilizza l'applicazione. In tal caso l'accesso ai dati è consentito dalle autorizzazioni:

- **pages_read_engagement**: consente di leggere contenuti (post, foto, video, eventi) pubblicati dalla Pagina, leggere i dati dei follower (inclusi nome, PSID) e immagine del profilo e leggere metadati e altri dati statistici sulla Pagina;
- **pages_read_user_content**: permette di leggere i contenuti generati dagli utenti sulla Pagina, come post, commenti e valutazioni degli utenti o di altre Pagine, e di eliminare i commenti degli utenti sui post della Pagina. Inoltre, consente alla tua app di leggere i post in cui la Pagina è citata.

2.2.4 Rate limit

Tutte le chiamate che vengono effettuate all'API Graph vengono conteggiate al fine di limitare il numero di chiamate che vengono effettuate in un determinato periodo temporale.

Il rate limit [19] è suddiviso in due categorie principali:

1) platform Rate Limit: si applicano alle chiamate relative all'*API Graph* o all'*Instagram Basic Display API*. Nel caso in cui la chiamata è effettuata con l'uso di un token d'accesso dell'app, il numero di chiamate totali che si possono effettuare in 1h è calcolato moltiplicando il numero di utenti che utilizza l'app per una costante pari a 200.

Al contrario se venisse utilizzato un token d'accesso utente, il conteggio delle chiamate è calcolato come il numero di chiamate che un utente può effettuare in 1h. Per questioni di privacy Facebook non condivide tale dato.

2) business Use Case (BUC) Rate Limits: in questo caso rientrano tutte le chiamate effettuate all'API marketing o all'API delle pagine (pages API) attraverso un token d'accesso della pagina. I limiti che vengono applicati dipendono dagli endpoint che si vogliono interrogare.

Nel caso di query alle pages API verrà applicato il platform Rate Limit o il BUC Rate Limits a seconda del tipo di token utilizzato; in caso di application o user access token verrà considerato il primo caso mentre se si utilizza un page access o system user access token⁴ si considererà il secondo caso.

⁴ Tipo di token appartenente al Business Management APIs [21], non trattato in tale sede.
Anno Accademico 2020/2021

Nel caso specifico dell'applicazione da me sviluppata vengono applicati i limiti descritti al punto 1), in particolare quelli relativi all'utilizzo di un user access token. Bisogna evidenziare che con le API 2.8 non si è mai incorsi in nessuna limitazione anche considerando il fatto che il massimo numero di chiamate non era noto e veniva applicato a discrezione della piattaforma.

Nel capitolo relativo al test illustrerò i risultati ottenuti dall'applicazione nella sua versione originaria e le differenze con l'ultima versione (11.0).

È possibile anticipare che il software, dopo l'aggiornamento, risulta essere pienamente funzionale entro i limiti dettati dall'utilizzo di un account non aziendale con la quale vengono svolti tali test.

3. Obiettivi software

Come annunciato nei capitoli precedenti il nostro obiettivo primario, nell'ambito della *Sentiment Analysis*, riguarda l'acquisizione dei dati, dal social network Facebook, riguardanti il mondo della moda. Per adempiere a tale obiettivo si è reso necessario creare un software che permettesse l'acquisizione dei dati da un numero predefinito di pagine Facebook e successivamente, dopo averle analizzate, trovarne di nuove.

Il software svolge, per le motivazioni sopracitate, dei compiti che possono raggrupparsi in due gruppi:

- monitoraggio pagine;
- discovery pagine.

L'applicazione è stata realizzata in Java utilizzando la libreria open-source Facebook4J [20] che permette una facile implementazione delle API di Facebook. Tale scelta è stata dettata dal fatto che l'applicazione doveva poter essere eseguita su differenti sistemi operativi indipendentemente dalla piattaforma di sviluppo. Per quanto riguarda il salvataggio delle informazioni è stato scelto un database MySQL parte del pacchetto software di sviluppo WAMPP.

3.1 Monitoraggio pagine

Il compito di monitoraggio delle pagine consiste, partendo da una lista predefinita di pagine (Appendice - 70 -A2) riguardanti il mondo della moda, di acquisire un determinato numero iniziale di post al fine di poterli analizzare.

Gli elementi fondamentali di un post (paragrafo 1.1.1) che vengono salvati durante la fase di acquisizione sono:

- testo del post;
- commenti;

- numero di like;
- hashtags;
- pagine taggate.

Prendendo in esempio il post analizzato (Figura 3-1) nel paragrafo 1.1.1 i dati che andremo ad ottenere saranno:



Figura 3-1: Esempio di post su Facebook [<https://www.facebook.com/nike/posts/10155790703238445>]

- Testo del post: “Outplay yourself. [Roger Federer](#), Melbourne's defending champion and holder of 19 major titles, just beat his own record, by winning 20. [#justdoit](#)”;
- Hashtag: [#justdoit](#);

- Tag: [Roger Federer](#);
- Like: 48368.

Se le pagine non sono mai state monitorate in precedenza (es: primo avvio dell'applicazione o inserimento di una nuova pagina) verranno presi i primi 50 post (valore modificabile) partendo dall'ultimo pubblicato dalla pagina.

Al contrario se sono presenti post per la pagina considerata, verranno presi solamente quelli compresi tra la data dell'aggiornamento e la data dell'ultimo post salvato.

Il limite dei 50 post è stato inserito sia per velocizzare il processo di grabbing e sia per evitare eventuali limiti imposti alle chiamate (2.2.4) dalla piattaforma.

Per ogni post, inoltre, vengono presi i primi 10 commenti (valore modificabile) se il post risulta essere stato appena acquisito, mentre se esistente (aggiornamento) il limite è pari a 30. Così come per i post gli elementi che vengono estratti dai commenti sono:

- testo del commento;
- hashtag;
- pagine taggate.

In sintesi, con l'avvio del programma verranno effettuate le operazioni di grabbing andando a prendere per ciascuna pagina post e relativi commenti, estrapolando i dati, sopracitati, di nostro interesse.

3.2 Discovery pagine

Dopo la fase di grabbing dei dati si ha la fase di discovery delle pagine. Tale fase consiste nell'analizzare i dati ottenuti nella fase precedente e ricavare da quest'ultimi nuove pagine che potrebbero risultare collegate, nel nostro caso, al mondo della moda.

Al primo avvio del programma verranno presi (con i valori di default) i primi 50 post di ogni pagina e per ogni post i primi 10 commenti; da tali post e commenti otterremo un certo numero di tag ad altre pagine da cui andremo ad acquisire, nuovamente, post e commenti (in questa fase limitati rispettivamente a 5 e 10) al fine di ottenere nuove pagine taggate.

È possibile indicare, nei parametri dell'applicazione, il numero di iterazioni che il software deve eseguire per ripetere le operazioni sopradescritte.

3.2.1 Esempio discovery pagine

Consideriamo di aver acquisito i post e i commenti di una determinata pagina che l'utente ha scelto di monitorare (es: [@gucci](#)) e che i risultati ottenuti nella fase di monitoraggio siano:

- pagine: [@vogueitalia](#), [@ellesse](#), [@gazzettadelloSPORT](#);
- hashtag: [#italia](#), [#roma](#), [#jumper](#), [#jacket](#), [#sneakersitalia](#).

Non sappiamo se le pagine ottenute sono relative o meno al mondo della moda e quindi l'utente decide di avviare la fase di discovery sulla base dei risultati ottenuti in precedenza per un numero totale di due iterazioni.

Durante la prima iterazione l'applicazione ottiene i post e i relativi commenti delle pagine [@vogueitalia](#), [@ellesse](#), [@gazzettadelloSPORT](#) e per ognuna di essa verrà analizzato il testo di ogni singolo post, per determinare se la pagina tratta di argomenti legati o meno al mondo della moda.

Se il numero di post che riguardano argomenti del mondo della moda supera, in percentuale, una determinata soglia (configurabile dall'utente e nel nostro caso corrispondente al 60% dei post) la pagina risulta essere idonea e, per tale motivo, verrà considerata per la prossima fase di acquisizione dati.

Dopo tali operazioni le pagine idonee risulteranno essere [@vogueitalia](#) e [@ellesse](#).

A questo punto verrà eseguita la seconda iterazione che considererà le pagine e gli hashtag trovati in precedenza per le pagine [@vogueitalia](#) e [@ellesse](#):

- pagine: [@naomicampbell](#), [@britneyspears](#), [@louisvuitton](#);
- hashtag: [#photovoguefestival](#), [#concertoroma](#), [#LVMenSS22](#).

Le pagine che risulteranno idonee dopo l'applicazione del filtro saranno [@naomicampbell](#) e [@louisvuitton](#).

Allo stesso tempo vengono sottoposti al filtro anche gli hashtag trovati nelle due iterazioni escludendo quelli che non fanno parte del mondo della moda; nel caso di esempio verranno considerati validi gli hashtag [#jumper](#), [#jacket](#), [#sneakersitalia](#), [#photovoguefestival](#), [#LVMenSS22](#).

Gli hashtag considerati idonei verranno utilizzati per la ricerca di utenti Twitter poiché la versione 2.8 delle API di Facebook non permetteva la ricerca tramite hashtag se non previa autorizzazione.

Partendo dalla pagina iniziale [@gucci](#), grazie al meccanismo di discovery e filtraggio, abbiamo scoperto altre quattro pagine d'interesse: [@vogueitalia](#), [@ellesse](#), [@naomicampbell](#) e [@louisvuitton](#); tali pagine verranno segnalate all'utente, tramite un report via e-mail, che deciderà o meno di aggiungerle alla lista delle pagine da monitorare.

4. Struttura del software

Dopo una breve descrizione sul funzionamento del software in questo capitolo andremo a descriverne il funzionamento, le funzionalità e i dati da esso utilizzati. Come abbiamo appena visto il software è strutturato in due fasi principali:

- 1) fase di grabbing: consiste nell'acquisizione dati dalle pagine sottoposte a monitoraggio;
- 2) fase di discovery e filtraggio: i dati ottenuti nel punto precedente vengono analizzati in modo da poter ricavare nuove pagine da monitorare;

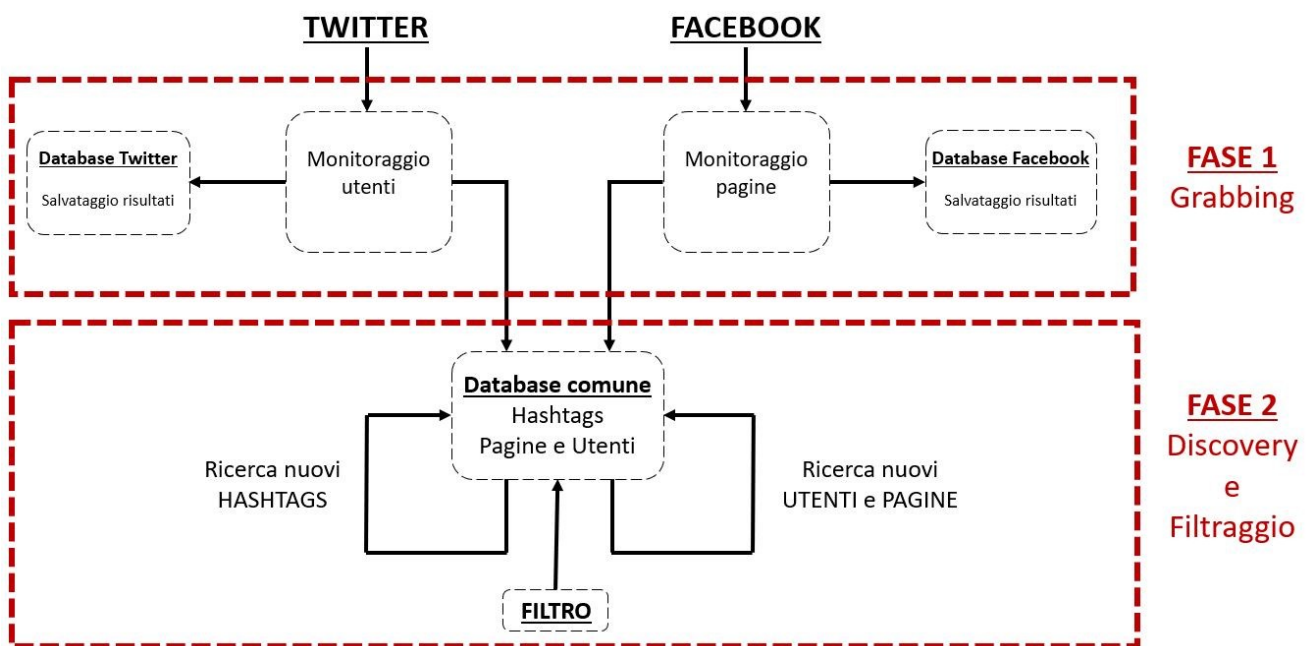


Figura 4-1: Struttura del software [Andrea Lagone, "Ricerca ed analisi di tweet per il mondo della moda"]

In Figura 4-1 è mostrata una visione generale del funzionamento del software ottenuta integrando le operazioni basilari, nelle varie fasi, per i due social network oggetti nel nostro studio. Notiamo che ogni applicativo utilizza un database dedicato (*Database Twitter* e *Database Facebook*) per il salvataggio delle informazioni relative alla propria piattaforma e un database in comune (*Database Comune*) utilizzato per i dati che verranno utilizzati nella fase di discovery.

4.1 Fase di grabbing

La prima operazione che l'utente deve effettuare è quella di fornire, a entrambi gli applicativi, una lista di pagine (facebook) e utenti (twitter) di partenza in modo da poter avviare la fase di grabbing dei dati (diagramma di flusso in Figura 4-4); l'utente può fornire tali informazioni tramite riga di comando e a tal riguardo le operazioni possibili per le pagine sono:

- a. aggiunta pagina tramite nome o link;
- b. cancellare una pagina
- c. disattivare o attivare il monitoraggio di una pagina;
- d. inserire una pagina trovata nella fase di discovery all'interno della lista delle pagine da monitorare; ciò comporta il trasferimento delle informazioni dal *Database Comune* al *Database Facebook*;
- e. avviare la fase grabbing e/o aggiornamento delle pagine da monitorare;
- f. avviare la fase di discovery.

Una volta che sono presenti delle pagine da monitorare, l'utente attraverso l'opzione "e", può avviare la routine che si occupa, per ogni pagina, del grabbing dei post; al fine di comprendere meglio come viene realizzata tale fase, nel paragrafo che segue verrà descritta la procedura di acquisizione dati da una pagina.

4.1.1 Acquisizione dati

Preso una pagina e avviata la procedura di acquisizione (Figura 4-2), la prima operazione consiste nell'imporre i limiti dei post e commenti che devono essere scaricati; per fare questo bisogna controllare se la pagina è inserita nel *Database Facebook* o nel *Database Comune*. Quest'ultimo caso si verifica quando la procedura di acquisizione dati è avviata dalla fase di discovery (Fase 2) e quindi si rende necessario limitare, in misura maggiore, il numero di post e commenti poiché il numero di pagine che si andranno ad analizzare è mediamente maggiore del numero di pagine inserite dall'utente.

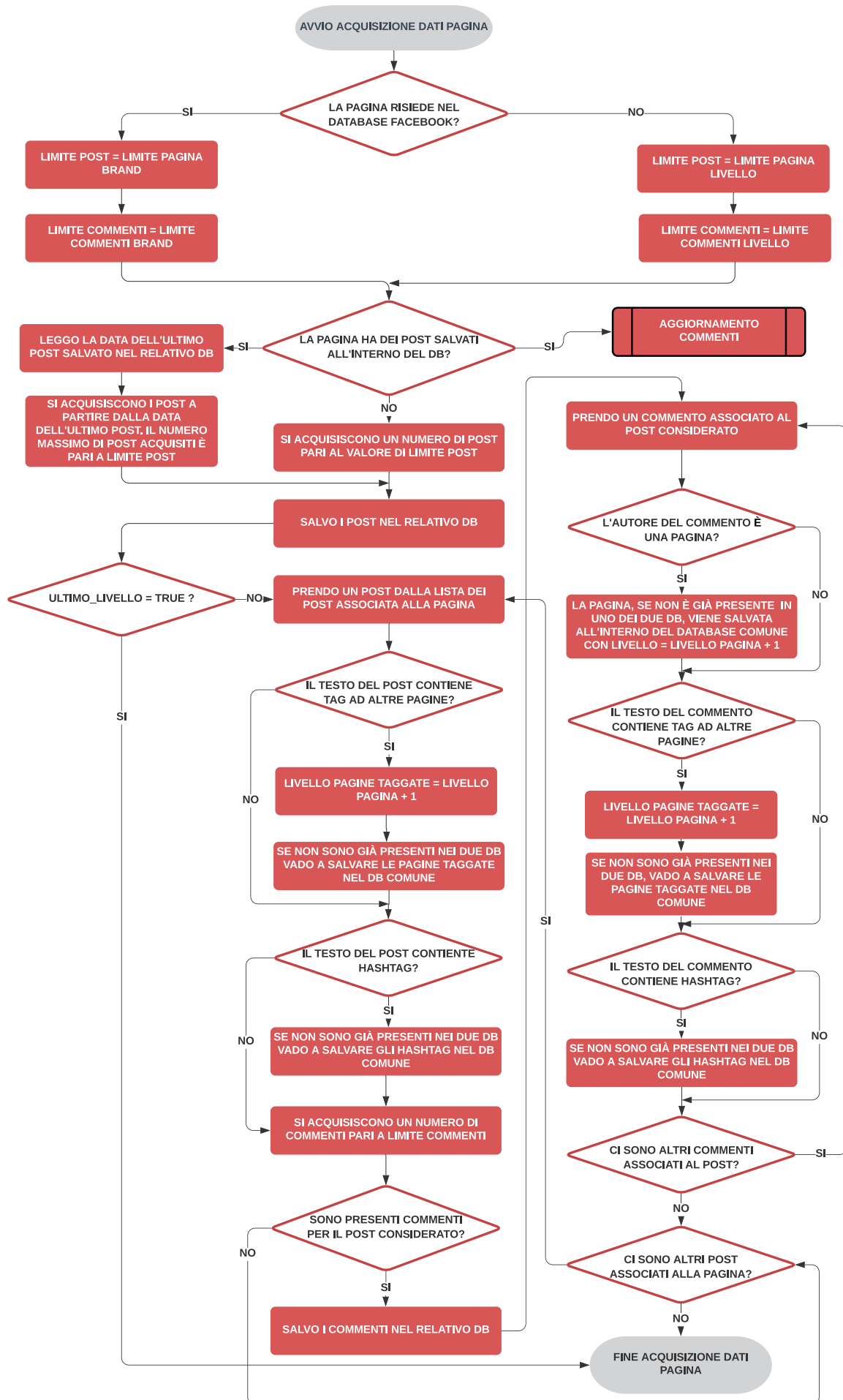


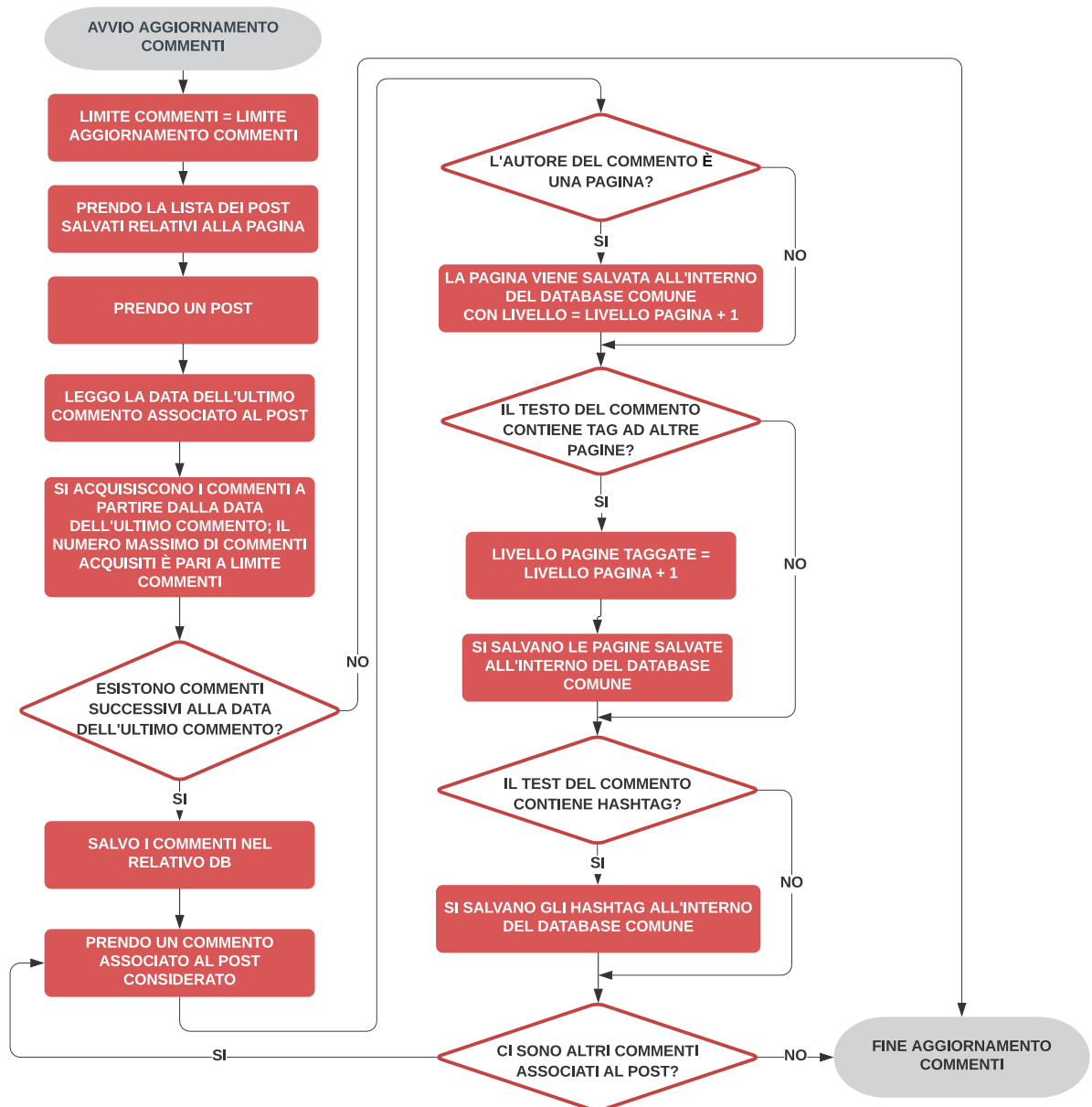
Figura 4-2:Diagramma di flusso della procedura di acquisizione dati

I limiti relativi a post e commenti sono editabili all'interno del file di configurazione dell'applicazione e sono complessivamente cinque:

- 1) numero_post_pagina_brand: numero massimo di post da scaricare per pagina da monitorare (salvata in *Database Facebook*, inserita dall'utente o una ammessa dopo la fase di discovery);
- 2) numero_post_pagina_livello: numero massimo di post da scaricare, nella Fase 2, per pagina scoperta nella Fase 1 (salvata in *Database Comune*);
- 3) limite_commenti_post_brand: numero massimo di commenti da scaricare per ogni post relativo a una pagina da monitorare;
- 4) limite_commenti_post_discovery: numero massimo di commenti da scaricare, nella Fase 2, per ogni post relativo a una pagina scoperta nella Fase 1;
- 5) limite_commenti_aggiornamento: numero massimo di commenti da scaricare nella routine aggiornamento commenti.

Successivamente se la pagina presenta dei post salvati all'interno del database vengono effettuate due operazioni:

1) Aggiornamento commenti (



): viene avviata la routine per aggiornare i commenti di un post salvato nel DB. Tale routine viene eseguita sulla pagina interessata dalla quale si prendono la lista dei post ad essa associata.

Per ogni post all'interno della lista viene presa la data del commento più recente salvato e successivamente viene effettuata una chiamata per ottenere i commenti del post a partire da tale data. Se non esistono post più recenti la routine viene terminata. Al contrario, se sono presenti dei commenti essi vengono salvati nel relativo database. Successivamente si avvia, per ogni nuovo commento, la procedura (descritta successivamente) per ricavare hashtag e pagine taggate.

2) Update post mancanti: si ottiene la data dell'ultimo post salvato nel database e si effettua una chiamata per scaricare i post pubblicati dopo tale data.

Al contrario, se la pagina non presenta nessun post si acquisiscono i post in base al numero fissato dal rispettivo limite.

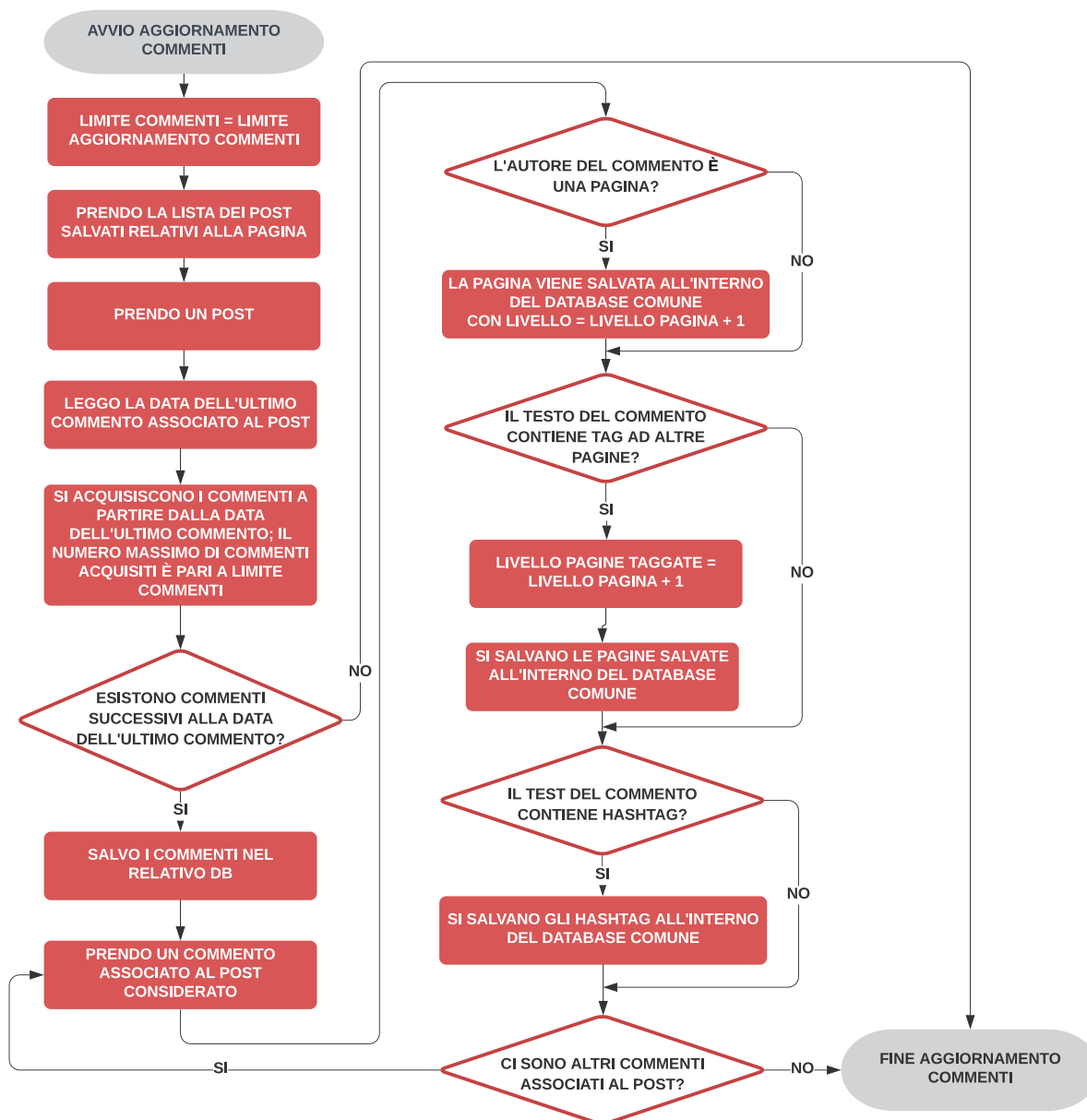


Figura 4-3:Diagramma di flusso relativo alla routine di aggiornamento commenti

Per ogni post si andranno a memorizzare all'interno del *Database Facebook* o del *Database Comune* i seguenti elementi:

- testo del post;
- numero di like;
- testo dei commenti.

Le pagine che vengono taggate direttamente nel testo del post sono contrassegnate come importanti (solo se la pagina da cui si acquisiscono i dati è contenuto in *Database Facebook*) le sia per differenziarli da eventuali tag presenti nei commenti, e sia perché potrebbero essere interessanti poiché riguardanti tematiche legate al mondo della moda; tali pagine vengono comunque sottoposte al filtro nella Fase 2 e, nel caso risultassero idonee, sono contraddistinte come importanti nel report per l'utente.

Per quanto riguarda gli hashtag ricordiamo che per loro definizione sono aggregatori tematici e riguardano perciò argomenti ben specifici, per tale motivo se l'hashtag è contenuto nel testo del post (**hashtag brand**) esso è contrassegnato come importante e viene posto nella lista degli hashtag da monitorare all'interno del *Database Comune*; inoltre, essi vanno a popolare il *dizionario hashtag brand* utilizzato dal Filtro Hashtag (paragrafo 4.3).

In modo analogo andiamo ad analizzare il testo dei commenti associati al post. A differenza del caso precedente gli autori dei commenti possono essere pagine o utenti; nel primo caso le pagine potrebbero essere collegate al mondo della moda (es: un negozio di abbigliamento che commenta un post di Armani) e contenere a sua volta informazioni interessanti per l'utente. Per tale motivo, se l'autore del commento è una pagina, essa viene salvata all'interno del *Database Comune* al pari delle pagine taggate.

Analogamente alle pagine taggate nel testo del post, anche quelle presenti nei commenti vengono salvati all'interno del *Database Comune* con l'unica differenza che non vengono contrassegnati come importanti. Gli hashtag dei commenti (**hashtag livello**) verranno salvati in una lista differente per distinguerli da quelli presenti nel testo del post.

In definitiva verranno memorizzati solamente gli hashtag e le pagine che non siano già presenti nei rispettivi database (es: una pagina scoperta in precedenza che non ha passato la fase di filtraggio); inoltre, tutte le pagine taggate e gli hashtag (tranne gli

hashtag brand) sono abilitati al controllo ma esclusi dal monitoraggio; quest'ultimo può essere attivato nella Fase 2 (dopo il filtraggio) oppure manualmente dall'utente.

La fase di monitoraggio degli hashtag importanti, per le motivazioni riportate nel capitolo precedente, viene effettuata esclusivamente con l'applicazione riguardante Twitter, di conseguenza è consigliabile avviare, a parità di pagine/utenti seguiti nelle due piattaforme, prima il monitoraggio con Facebook, ottenendo così un certo numero di hashtag da monitorare, e successivamente con Twitter in modo da acquisire un maggior numero di utenti nella Fase 2.

4.1.2 Fase di grabbing (Fase 2)



Figura 4-4: Diagramma di flusso relativa alla fase di grabbing

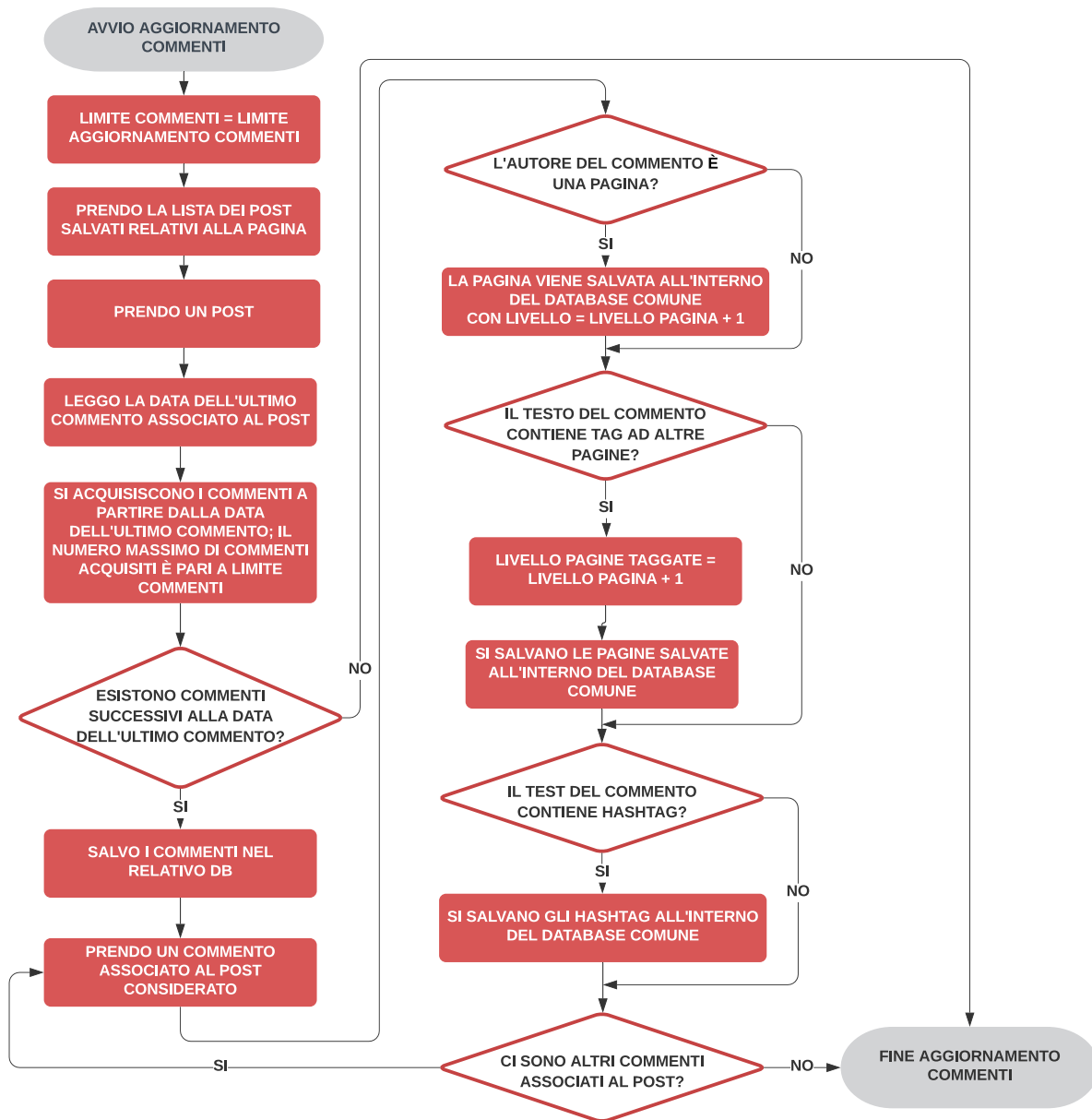
La fase di grabbing (Figura 4-4) consiste nell'acquisire i dati di nostro interesse a partire dalle pagine presenti nel *Database Facebook*. Tali pagine possono essere state
Anno Accademico 2020/2021

appena inserite (dall'utente o dalla fase di discovery) o essere già presenti nel database, in tal caso verrà effettuato l'update dei post mancanti e l'aggiornamento dei commenti riguardante i post già presenti nel database.

Dal momento che l'utente può abilitare o disabilitare il monitoraggio di una pagina, dopo l'avvio della fase di grabbing, si andranno a prendere unicamente le pagine abilitate all'interno del *Database Facebook*.

Dopo aver ottenuto la lista delle pagine da monitorare per ognuna di esse vado a richiamare la funzione di acquisizione dati (illustrata nel paragrafo precedente) così da ottenere hashtag e pagine taggate che vengono salvati all'interno del *Database Comune*.

Prima di procedere alla descrizione della fase di discovery, bisogna introdurre il concetto di livello associato a ogni pagina; come abbiamo notato nei diagrammi di flusso di Figura 4-2 e



, le pagine salvate nel *Database Comune* presentano una proprietà chiamata **livello**; quest'ultimo indica a che iterazione è stata trovata una determinata pagina.

Tutte le pagine presenti nel *Database Facebook* hanno un livello pari a zero. Una volta avviata la Fase 1, le pagine taggate nel testo del post o nei commenti verranno salvati, nel *Database Comune*, con un livello pari a uno e ammesse al controllo nella Fase 2.

4.2 Fase di discovery

Al termine della fase di grabbing abbiamo ottenuto un certo numero di pagine di livello 1 salvate all'interno del *Database Comune*; tali pagine non presentano post

salvati all'interno del database per cui non è possibile sapere a priori se sono legate o meno al mondo della moda.

La fase di discovery ci permette di determinare quali pagine di livello 1 potrebbero essere interessanti per l'utente, e partendo da quest'ultime ricavare ulteriori pagine di livello superiore.

Consideriamo che ,attraverso il comando "F", l'utente abbia avviato la fase di discovery specificando un numero "n" (con $n > 1$) di iterazioni; durante la prima iterazione ($n=1$) vengono prese tutte le pagine di livello 1 dal *Database Comune*, abilitate al controllo, e per ognuna di essa si avvia la procedura di acquisizione dati (4.1.1) descritta in precedenza per la Fase 1. Tutte le pagine e gli hashtag trovati in tale procedura sono salvati con un livello pari a 2 ($\text{livello} = n + 1$) e ciò indica che si sono effettuate due iterazioni, rispetto alle pagine presenti in *Database Facebook* ($\text{livello}=0$), per trovare tali pagine.

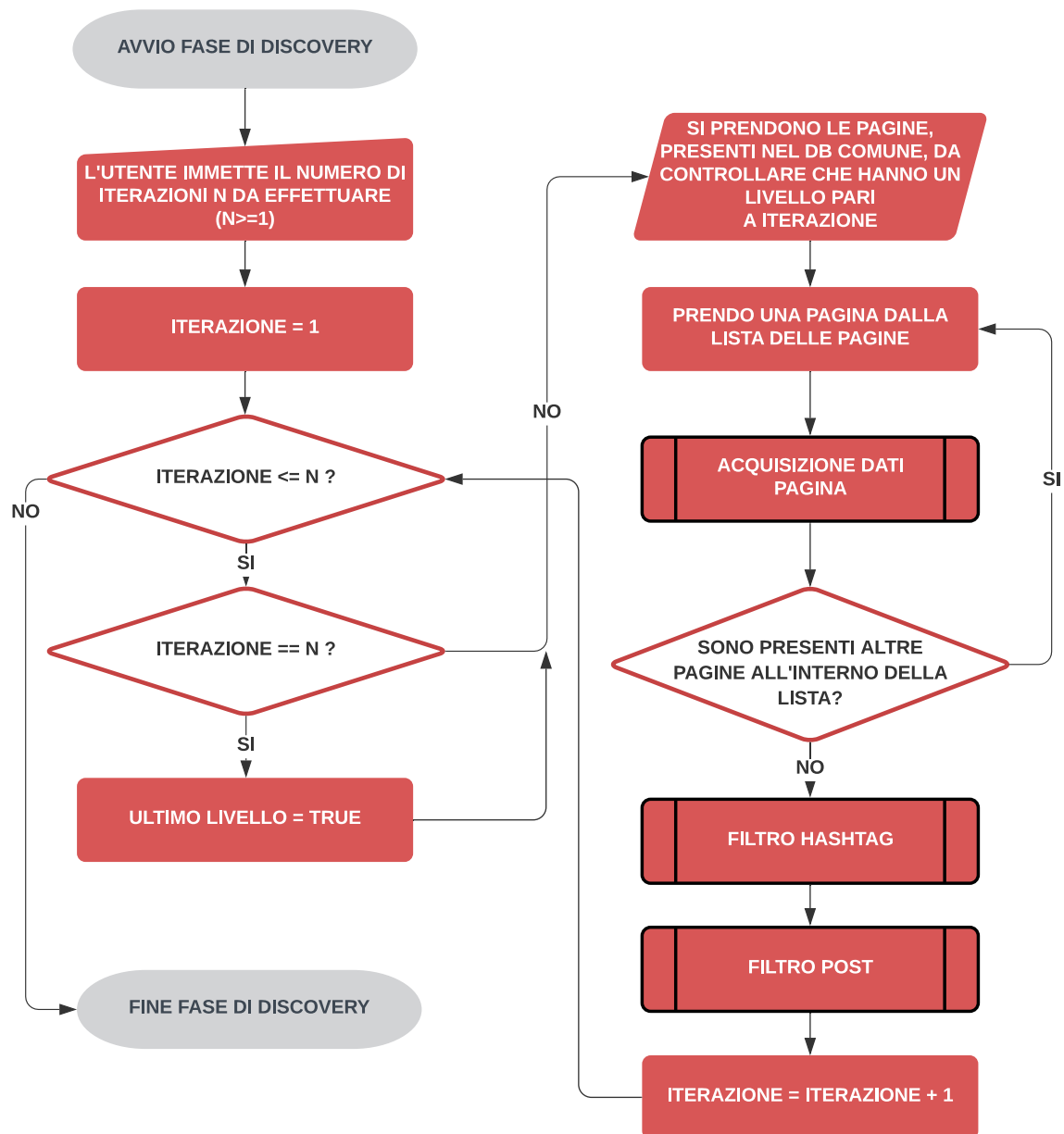


Figura 4-5: Diagramma di flusso della fase di discovery (Fase 2)

La procedura continua fino a quando non sono presenti altre pagine di livello 1 e, in seguito, verrà applicato il filtro a tutti gli hashtag e pagine di tale livello; nel caso in cui risultassero idonei verranno contrassegnati come “ammessi” al monitoraggio. Successivamente verrà incrementato il valore dell’iterazione di un’unità; nel nostro caso “iterazione” assumerà il valore 2 e l’intero ciclo verrà ripetuto considerando questa volta le pagine di livello 2 che sono state scoperte nell’iterazione precedente. Una volta raggiunta l’iterazione n-esima il valore della variabile booleana “ultimo livello” assumerà il valore “true”; tale valore indicherà alla procedura di acquisizione dati, di non analizzare il testo del post e i commenti per la ricerca di pagine taggate o

hashtag. Infatti, nell'ultima iterazione i post verranno ugualmente acquisiti ma ai soli fini dell'applicazione del Filtro post livello.

4.3 Filtro

Il filtro è la parte del software che si occupa di esaminare il testo dei post e gli hashtag al fine di abilitarli per il monitoraggio. Esso viene richiamato dalla fase di discovery una volta terminata l'acquisizione dati di tutte le pagine di un determinato livello.

Il filtro si basa su tre dizionari di default salvati all'interno del *Database Comune*:

- 1) *dizionario brand*: è costituito da marchi di moda esistenti (es: Gucci, Armani, Loro Piana, Nike, Adidas, ecc.);
- 2) *dizionario moda*: è costituito dai termini che descrivono le diverse tipologie di articoli di moda e abbigliamento (es: a macchie, a maglia larga, a costa inglese, chiusura, cinturino alla caviglia, ecc.) in italiano e in inglese.
- 3) *dizionario hashtag brand*: realizzato in maniera dinamica tramite gli hashtag identificati all'interno dei post delle pagine da monitorare (pagine di livello=0); inoltre, tali hashtag vengono utilizzati per la ricerca di nuovi utenti di interesse all'interno del social network Twitter.

Il software permette di specificare quattro parametri relativi al filtro:

- 1) *soglia_validità_post*: numero minimo di parole riguardanti la moda (*dizionario moda*) all'interno del testo di un post affinché risulti valido;
- 2) *lunghezza_minima_parola*: lunghezza minima di una parola affinché venga confrontata con le parole all'interno del dizionario moda;
- 3) *lunghezza_minima_hashtag*: rappresenta la lunghezza minima di un hashtag. Se l'hashtag ha una lunghezza maggiore di quella specificata, allora il suo

contenuto verrà confrontato con le parole presenti all'interno dei dizionari brand, moda e Hashtag Brand;

- 4) *soglia_minima_post*: percentuale minima di post, rispetto al totale salvato nel database, che una pagina deve avere affinché venga considerata importante.

Il filtro è composto da due elementi principali: filtro hashtag e filtro post livello. Il primo è applicato per determinare quali hashtag sono ammessi al monitoraggio da parte dell'applicazione realizzata per Twitter, mentre il secondo è utilizzato per determinare se una pagina può risultare interessante per l'utente in base al contenuto dei suoi post.

4.3.1 Filtro Hashtag

La prima fase di filtraggio consiste nell'esaminare tutti gli hashtag relativi ad uno specifico livello. Una volta ottenuta, tramite un'interrogazione al *Database Comune*, la lista degli hashtag, per ogni elemento della lista si verifica se la lunghezza dell'hashtag è minore o uguale al valore del parametro lunghezza_minima_hashtag; in caso affermativo, l'hashtag potrebbe essere il nome di un brand e per tale motivo, ogni elemento del *dizionario brand* viene confrontato con l'hashtag in questione. Nel caso in cui un elemento del *dizionario brand* risultasse uguale all'hashtag verrà abilitato il monitoraggio per l'hashtag in questione, in caso contrario verrà disabilitato.

Se la lunghezza dell'hashtag dovesse superare il valore del parametro lunghezza_minima_hashtag si verificherà che l'hashtag non è contenuto all'interno di un elemento di uno dei tre dizionari; se ciò dovesse accadere, anche per un solo elemento, l'hashtag verrà ritenuto valido.

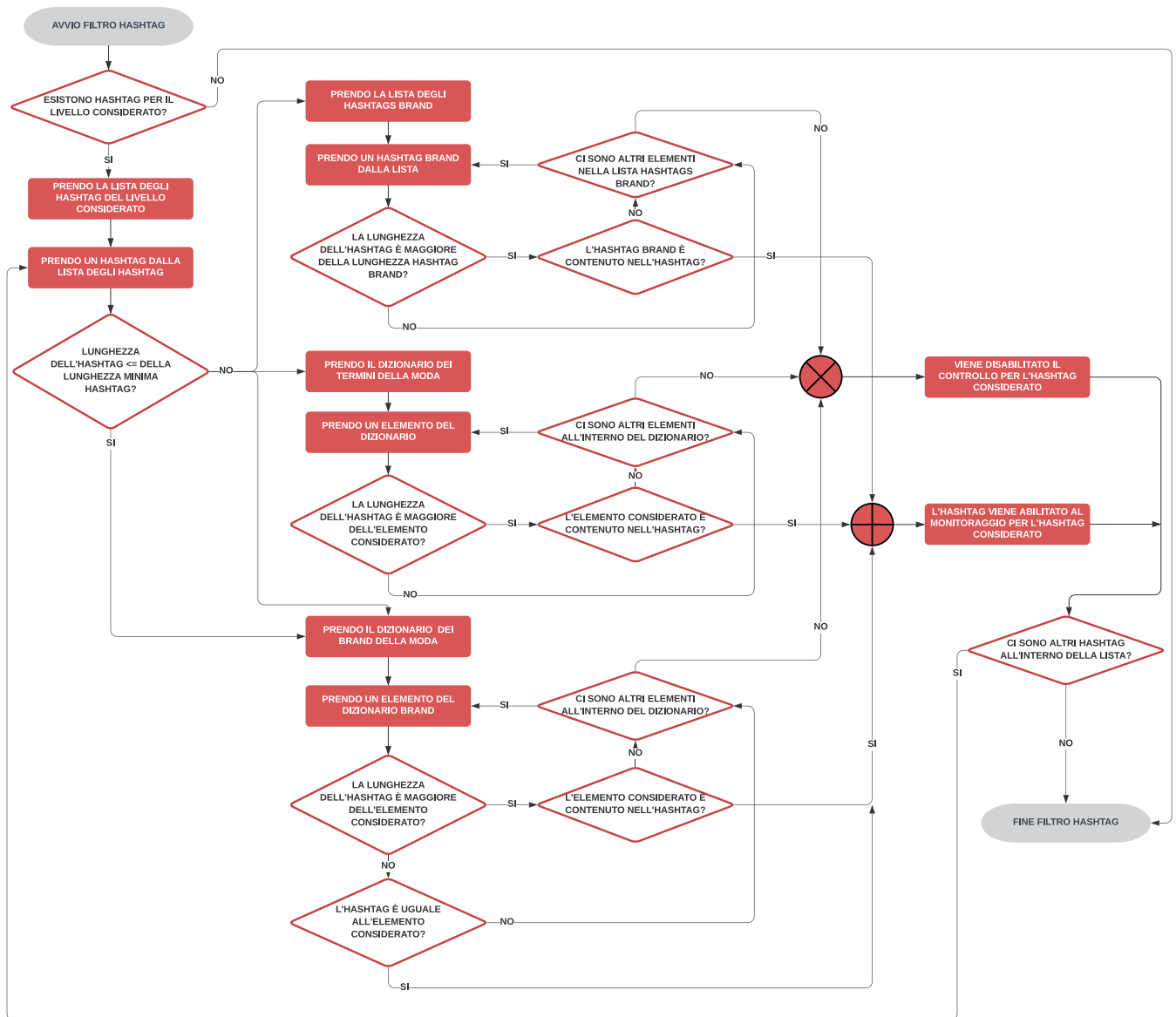


Figura 4-6: Diagramma di flusso del Filtro Hashtag

4.3.2 Filtro post livello

Una volta terminato il filtraggio degli hashtag si procede all'analisi dei post delle pagine di un determinato livello al fine di decidere se una pagina può essere considerata ammissibile per la fase di monitoraggio.

Per ogni pagina presente all'interno della lista delle pagine si va a prendere la lista dei post associati; all'inizio della routine, attraverso le variabili `post_validi` e `parole_moda`, si tiene traccia del numero di post idonei di una determinata pagina e del numero di parole associate al mondo della moda per un dato post.

Entrambe le variabili vengono poste pari a zero quando si considera una nuova pagina o un nuovo post.

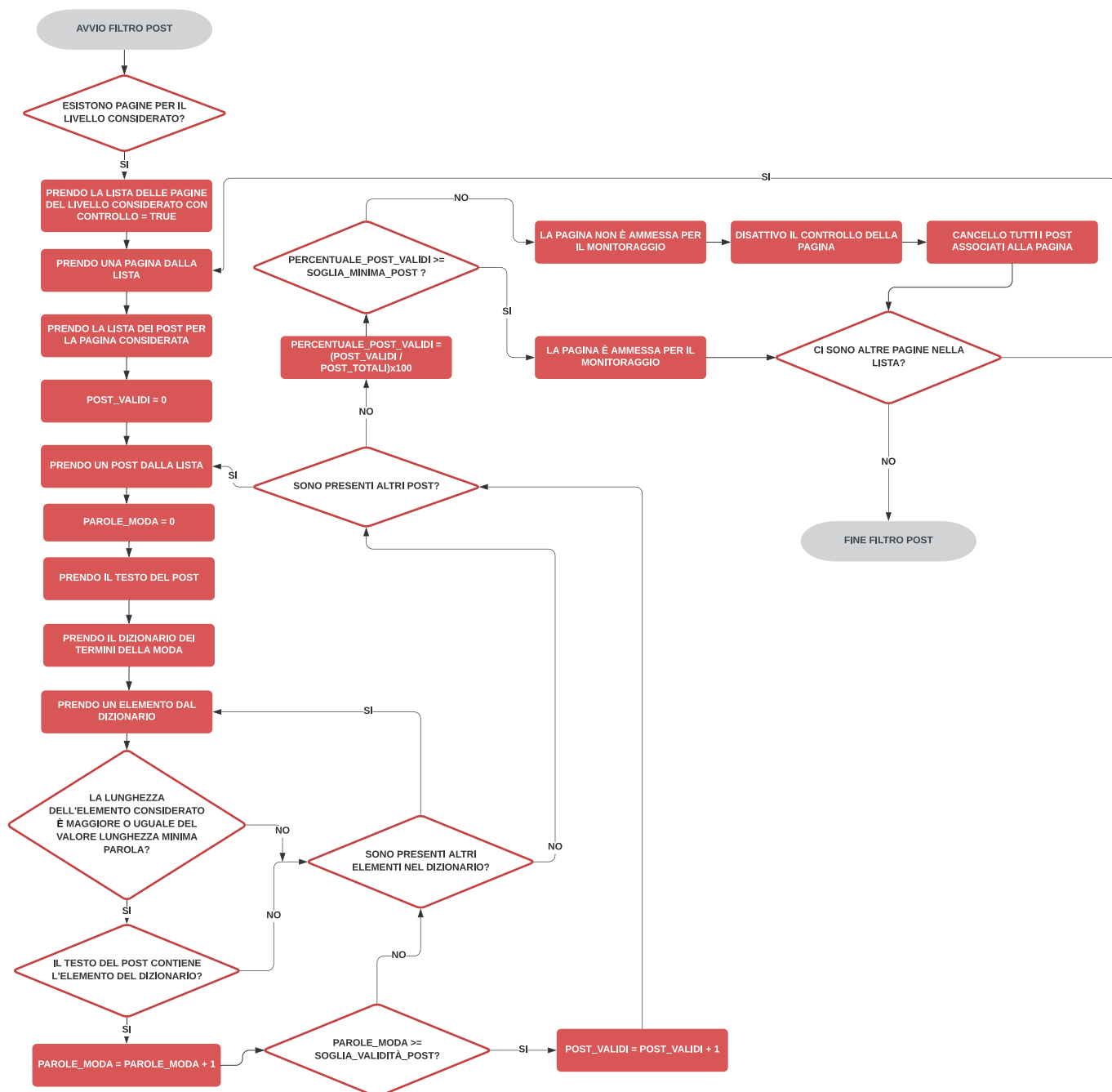


Figura 4-7: Diagramma di flusso relativo al Filtro post livello

Si presume che la pagina pubblicizzi i suoi prodotti, tramite i post, descrivendoli con uno o più termini presenti all'interno del *dizionario moda*, pertanto viene valutato se il post contiene uno o più elementi di tale dizionario; per velocizzare l'analisi verranno considerati solo gli elementi moda che presentano una lunghezza maggiore del valore "lunghezza_minima_parola".

Per ciascun termine trovato all'interno del testo verrà incrementato il valore della variabile "parole_moda" e se quest'ultimo ha un valore maggiore o uguale del parametro "soglia_validità_post" il post viene considerato valido terminando il ciclo di confronto; se ciò non dovesse accadere si proseguirebbe fino a quando non ci sono più termini da valutare in *dizionario moda*. Se il post è considerato idoneo viene incrementato il numero di post validi e, se presente, si passa al post successivo. Una volta concluso il controllo si calcola la percentuale di post validi, rispetto al totale, e se tale risultato è maggiore (o uguale) al valore del parametro "soglia_minima_post" la pagina risulta valida per il monitoraggio. Diversamente il controllo della pagina, in fase di discovery, viene disabilitato e vengono cancellati tutti i post ad esso relativi presenti nel *Database Comune*. Mantenere la pagina, ma non i suoi post, all'interno del database permette di evitare una nuova fase di acquisizione dati nel caso in cui la pagina dovesse risultare taggata all'interno di qualche post o commento di un'altra pagina.

5. Test

In quest'ultimo capitolo verranno esposti i risultati ottenuti dall'applicazione in seguito allo svolgimento di un test sulle pagine [@Armani](#) e [@Benetton](#).

Il test è stato effettuato a fine 2016 con la versione 2.8 delle API Graph su una distribuzione Linux Ubuntu Server 14.04.5 LTS.

I parametri utilizzati per la fase di acquisizione dati (pagina - 45 -) sono:

- numero_post_pagina_brand = 10;
- numero_post_paginalivello = 30;
- limite_commenti_post_brand = 10;
- limite_commenti_post_discovery: 30.

Rispetto i parametri di default si è reso necessario diminuire il numero di post acquisiti per le pagine monitorate e aumentare quello relativo alle pagine da analizzare nella fase di discovery e filtraggio; tale scelta è dovuta in primo luogo a velocizzare la fase di acquisizione dati e in secondo luogo di avere un numero sufficiente di post per l'applicazione del filtro.

Per quanto riguarda i parametri del filtro (pagina - 54 -) abbiamo:

- soglia_validità_post=2;
- lunghezza_minima_parola= 4;
- lunghezza_minima_hashtag= 6;
- soglia_minima_post= 60.

Fase 1 (Grabbing)

Partendo dalle pagine ufficiali di Armani ([@Armani](#)) e Benetton ([@Benetton](#)) otteniamo i risultati espressi nella Tabella 1.

Tabella 1: Risultati test Fase 1

Elemento	Dati acquisiti
post	20
Hashtag brand	17
Hashtag livello (livello 1)	44
pagine taggate (livello 1)	5

Qui di seguito troviamo le immagini raffiguranti i risultati ottenuti per ciascun elemento:

- Post

	testo_post	likes
ARMANI	Giorgio Armani lancia le #GABagStraps, una collezi...	138
	Chris Pratt covered up in a midnight blue cashmere...	2087
	Opt for a coat with standout appeal in an eye-catc...	1038
	Stay warm without compromising style in comfortabl...	1288
	In occasione della serata inaugurale al Teatro All...	543
	HOLIDAY SPORTS: Get a jump start on your gift shop...	788
	Get to know !!! (Chk Chk Chk), special guest perfo...	624
	Breeze through your winter in #EA7!	39
	Time to unplug. Our new #EAConnected hybrid smartw...	3784
	Spread some delicious holiday cheer with Armani/Do...	696
BENETTON	White, black and beige. Check out the warm woolen ...	8484
	The red seamless #TV31100 sweater from United Colo...	583
	Around the world, 1 in 3 women experience some for...	346
	United Colors of Benetton is launching its first c...	589
	Wool and foliage patterns: check out the new winte...	5083
	The exclusive #StellaJeanForBenetton capsule colle...	360
	Plus qu'un jour avant l'arrivée de #StellaJeanForB...	365
	Un armario lleno de color, tejidos ligeros, estamp...	756
	True love is all about the small things.	626
	Humans communicate in many different ways. With cl...	350

Figura 5-1: Risultati dei post acquisiti durante la Fase 1

- Hashtag brand:



hashtag
16days
armanialist
armanidolci
armanistars
benetton
benettonss17
clothesforhumans
ea7
eaconnected
emporioarmani
emporioarmanisounds
gabagstraps
gabugatti
orangetheworld
stellajeannotforbenetton
tv31100
undercolors

Figura 5-2: Risultati dei brand acquisiti dopo la Fase 1

- Hashtag livello

hashtag
1988massacre
adr
afp
alassio
awareness
benetton_japan
bestgiftever
blogger
bomdia
bomdomingo
borderlesscelebration
campaignforjoy
chrispratt
culturalcrossover
culturalpromotion
dior
dolcegabbana
drmartens
emmastone
erhicallyenvisioned
ethicalfashion
ethicallyenvisioned
fendi
fw16
geekchic

Figura 5-3: Risultati degli hashtag trovati nei commenti o nei post delle pagine di livello ≥ 1

- Pagine taggate

id	nome_pagina	livello
[REDACTED]	!!! (Chk Chk Chk)	1
[REDACTED]	Campaignforjoy.org	1
[REDACTED]	Stella Jean	1
[REDACTED]	Spinnaker Boutique	1
[REDACTED]	EL-style	1

Figura 5-4: Pagine taggate trovate dopo la Fase 1 (per privacy sono stati censurati i rispettivi id)

Fase 2 (Discovery e filtraggio)

Tale fase è stata avviata con un numero totale di iterazioni pari a tre e i risultati complessivi, per ciascuna iterazione (livello) sono mostrati nella Tabella 2.

Tabella 2: Risultati ottenuti dalla Fase 2

Livello	Nuove pagine (pre filtraggio)	Pagine ammesse (post filtraggio)	Nuovi hashtag (pre filtraggio)	Hashtag ammessi (post filtraggio)
1	5	2	44	31
2	22	7	122	96
3	54	14	345	244

Nella figura seguente sono elencate le pagine ammesse al monitoraggio

- [Frankel's Delicatessen](#)
- BE MY GUEST (προσκλητηρια - μπομπονιερερς)
- [The Regent](#)
- COFFEEDENTIAL
- [Yilian Canizares](#)
- The Love Song Bar
- TSW
- [KOURTI MAGDALENE wedding photography](#)
- Global Down [Syndrome Foundation](#)
- Stella Jean
- Salma Hayek [Pinault](#)
- Vogue Italia
- Spinnaker Boutique
- Isa e Chia
- [Yiannis Sotiropoulos Photography](#)
- [Jada Pinkett Smith](#)
- [Ios Club Cocktail Bar](#)
- The12events
- Jennifer Hudson
- Auditorium Parco della Musica - Roma
- [Dead Heavens](#)
- TWRP
- Weddings in [Greece Be My Guest](#)

Differenze del test con le ultime API

Al momento della redazione di tale tesi l'applicazione così come è stata inizialmente realizzata non permette di replicare i test appena descritti poiché, come anticipato nel paragrafo 2.2.3, l'accesso all'*API Pages Search* non risulta essere più disponibile con un account di tipo non business. Inoltre, si deve essere autorizzati all'uso della funzione **Page Public Content Access** per accedere alle informazioni presenti all'interno di pagine pubbliche. Per ottenere l'accesso a tale funzione bisogna possedere un account di tipo business e sottoporre l'applicazione alla procedura di *App Review*.

In aggiunta, la nostra applicazione necessiterebbe di una rivisitazione per come vengono salvati i dati: attualmente i post e i commenti sono identificabili, ovvero è possibile ricavare l'autore di un post o commento partendo dall'id di quest'ultimi. Invece una prerogativa per l'accesso alla **Page Public Content Access** è la totale assenza di riferimenti personali nei dati salvati all'interno del database e l'impossibilità di risalire agli autori del contenuto; nel nostro caso ciò non dovrebbe risultare un problema poiché ciò che ci interessa conoscere sono nuove pagine e non le informazioni personali degli utenti attivi su una determinata pagina.

Diversamente l'applicazione è risultata parzialmente funzionante nel caso l'utente fosse proprietario di una o più pagine. Ciò è stato reso possibile attraverso le autorizzazioni *pages_read_engagement* e *pages_read_user_content*, che possono esse utilizzate anche senza account business. Anche in questo caso l'accesso a tali permessi è limitato alla versione “**Standard**” in quanto la versione “**Advanced**” è disponibile solamente dopo il processo di *App Review*.

Nelle immagini seguenti sono mostrati i risultati della *Fase 1* eseguita su una pagina in proprio possesso, dove sono stati pubblicati dei post e commenti di prova con citazioni a pagine di proprietà e pagine pubbliche; i risultati sono in linea con la versione originaria dell'applicazione.

indirizzo	id_pagina	nome_pagina	id_ultimo_post	controllo
https://www.facebook.com/Prova-FB-102469758344933	102469758344933	Prova-FB-102469758344933	102469758344933_105784618013447	1

Figura 5-6: Pagina di prova utilizzata con la versione 11.0 dell'API Graph

id_pagina	id_post	testo_post	id_ultimo_commento	likes
102469758344933	102469758344933_102469961678246	post 1	102469961678246_103928574865718	1
102469758344933	102469758344933_102470681678174	post 2 - senza commenti	NULL	1
102469758344933	102469758344933_103556374902938	post 3 #italy wink127 café www.google.com	103556374902938_104641184794457	1
102469758344933	102469758344933_103560718235837	sensori tecnologici wink127 café #sensori andrea l...	103560718235837_104640428127866	1
102469758344933	102469758344933_104452351480007	nuova uscita wink127 café #update	104452351480007_104639934794582	1
102469758344933	102469758344933_105784618013447	emporio armani the best!	105784618013447_105800964678479	0

Figura 5-7: Post acquisiti nella *Fase 1*

id_post	id_commento	testo_commento
102469758344933_102469961678246	102469961678246_102470011678241	commento 1
102469758344933_102469961678246	102469961678246_102470031678239	commento 2
102469758344933_102469961678246	102469961678246_102470055011570	commento 3
102469758344933_102469961678246	102469961678246_102471115011464	prova commento
102469758344933_102469961678246	102469961678246_102472558344653	naruto & kurama prova
102469758344933_102469961678246	102469961678246_103259888265920	#italia prova hash
102469758344933_102469961678246	102469961678246_103928574865718	prova commento pagina esterna
102469758344933_103556374902938	103556374902938_104641184794457	top!!! #top #top2 @andrea.langone.1
102469758344933_103560718235837	103560718235837_104452518146657	omg!!! #stile
102469758344933_103560718235837	103560718235837_104640428127866	prova up 2 #francia
102469758344933_104452351480007	104452351480007_104452424813333	i like it! #love
102469758344933_104452351480007	104452351480007_104639934794582	pova update 1 #prova 1

Figura 5-8: Commenti acquisiti nella *Fase 1*

id	nome_pagina	id_ultimo_post	controllo	livello
2001295643464257	wink127 café	2001295643464257_2001502826776872	1	1
114392070318875	provafba	NULL	0	1
350853578324893	naruto & kurama	350853578324893_1652736628136575	0	1
██████████	emporio armani	NULL	0	1

Figura 5-9: Pagine taggate ricavate da post e commenti

Successivamente è stata testata la Fase 2 nella quale si sono verificati gli effetti delle limitazioni delle nuove API; in tale fase è stato possibile accedere solo ai post delle pagine in proprio possesso (le prime tre in Figura 5-9) mentre per la pagina esterna ([Emporio Armani](#)) non è stato possibile accedere a tali informazioni poiché non si hanno le autorizzazioni necessarie.

Dal momento che le pagine di livello 1, individuate nella Fase 1, non sono accessibili la fase di discovery verrà interrotta prematuramente rendendo di fatto impossibile l'applicazione del filtro per la determinazione delle pagine di potenziale interesse per l'utente. Quest'ultimo può tuttavia decidere di valutare manualmente tali pagine anche se esse non possono essere soggette a monitoraggio per le motivazioni sopra riportate. Viceversa, il filtro verrà applicato correttamente a tutti gli hashtag di livello 1 dal momento che non sono necessarie altre informazioni per eseguire l'analisi.

Conclusioni

Inizialmente l'obiettivo principale di questa tesi era la realizzazione di un software in grado di fornire informazioni, riguardo il sentiment dei consumatori, alle aziende operanti nel mondo della moda.

Tuttavia, ci si è resi conto che per la realizzazione di tale obiettivo era necessario disporre di una o più fonti di dati dalla quale attingere le informazioni richieste; sebbene il web sia costituito da moltissime fonti d'informazioni, la scelta più naturale nell'ambito del nostro progetto, vale a dire il “**Social media sentiment analysis for fashion world**”, è ricaduta sui social network, in particolare Facebook e Twitter. Per tale motivo, il lavoro descritto in questa tesi ha riguardato lo sviluppo di un'applicazione che consentisse di ottenere le informazioni necessarie dal social network Facebook. Tale funzione è stata ampliata attraverso un algoritmo che permettesse di scoprire, dalle informazioni fornite dalle pagine e dal riscontro degli utenti, potenziali marchi d'interesse, conosciuti o meno, in modo da ampliare la quantità di dati a propria disposizione.

Come abbiamo visto la versione iniziale del software l'applicazione adempiva perfettamente ai nostri obiettivi ma nel corso di questi anni, a causa di numerosi eventi, le informazioni che prima erano liberamente accessibili ora richiedono specifiche autorizzazioni fornite solamente ad account aziendali dopo complesse procedure di verifiche.

Dopo aver aggiornato le funzioni alle nuove API, aggiornando le parti deprecate, l'applicazione risulta essere ancora funzionale entro i limiti delle autorizzazioni a propria disposizione.

Per poter continuare ad utilizzare il software bisogna in primo luogo assicurarsi di non salvare informazioni personali relative agli utenti che interagiscono con le pagine da monitorare; in secondo luogo, dobbiamo disporre di un account aziendale, rappresentare un'azienda o una società, e in ultimo luogo, dimostrare a Facebook come vengono trattati i dati messi a disposizione.

Alternativamente l'account aziendale potrebbe non essere richiesto se l'utente desidera monitorare una o più pagine in suo possesso ma in questo caso non sarà possibile sfruttare al 100% la funzione di discovery.

Il software inoltre potrebbe essere migliorato realizzando un'interfaccia grafica, possibilmente tramite un'applicazione web, in modo da facilitare l'interazione con esso.

Inoltre, potrebbero essere introdotti algoritmi basati sull'intelligenza artificiale e l'apprendimento automatico per l'analisi del testo in modo migliorare la fase di filtraggio dei contenuti.

In aggiunta per velocizzare l'esecuzione di acquisizione dati, il software, dopo un'opportuna modifica, potrebbe essere eseguito in più istanze ognuna con un proprio access token in modo da aumentare il numero di chiamate a propria disposizione.

In definitiva possiamo affermare che il software potrebbe essere ancora utilizzato per gli obiettivi prefissati ma deve essere continuamente aggiornato per far fronte alle continue evoluzioni delle API e dei cambiamenti riguardanti la privacy degli utenti.

Appendici

A1. Tabella versioni API Facebook

Tabella 3: Versioni dell'API Graph [<https://developers.facebook.com/docs/graph-api/changelog/versions>]

Versione API	Data di rilascio	Disponibile fino a
11.0	8/06/2021	da definire
10.0	23/02/2021	8/06/2023
9.0	10/11/2020	23/02/2023
8.0	04/08/2020	10/11/2022
7.0	05/05/2020	04/08/2022
6.0	03/02/2020	05/05/2022
5.0	29/10/2019	03/02/2022
4.0	29/07/2019	02/11/2021
3.3	30/04/2019	03/08/2021
3.2	23/10/2018	13/08/2019
3.1	26/07/2018	14/05/2019
3.0	01/05/2018	01/02/2019
2.12	30/01/2018	07/08/2018
2.11	07/11/2017	07/08/2018
2.10	18/07/2017	08/05/2018
2.9	18/04/2017	06/11/2017
2.8	05/10/2016	26/07/2017
2.7	13/07/2016	25/04/2017
2.6	12/04/2016	05/10/2016
2.5	07/10/2015	13/07/2016
2.4	08/07/2015	11/04/2016
2.3	25/03/2015	08/10/2015
2.2	30/10/2014	08/07/2015
2.1	01/10/2014	11/03/2015

2.0	01/10/2014	11/03/2015
1.0	01/10/2014	11/03/2015

A2. Lista dei brand monitorati e relativi account Facebook

Nome Brand	Nome account Facebook
<i>A.P.C.</i>	APCOfficialpage
<i>Reem Acra</i>	ReemAcraNY
<i>Agnona</i>	Agnona.Official
<i>Akris</i>	AkrisOfficial
<i>Alexander McQueen (brand)</i>	AlexanderMcQueen
<i>Alfred Dunhill Ltd.</i>	AlfredDunhill
<i>Charlie Allen (designer)</i>	officialcharlieallen
<i>Ally Capellino</i>	allycapellino.co.uk
<i>André Laug</i>	AndréLaugHauteCouture
<i>Andrew Marc</i>	AndrewMarcLeather
<i>Anne Fontaine (brand)</i>	AnneFontaineInc
<i>Aquascutum</i>	aquascutum1851
<i>Armani</i>	ARMANI
<i>Aspinal of London</i>	aspinaloflondon
<i>Asprey</i>	Asprey
<i>Dominique Aurientis</i>	dominiqueaurintis
<i>Agnès b.</i>	agnesb.officiel
<i>Badgley Mischka</i>	BadgleyMischka
<i>Balenciaga</i>	Balenciaga
<i>Bally Shoe</i>	pages/Bally-Shoe/112257075453144
<i>Balmain (fashion house)</i>	balmainparis
<i>Neil Barrett (fashion designer)</i>	NEILBARRETT
<i>Belle & Bunty</i>	belleandbunty
<i>Berluti</i>	Berluti
<i>Beulah London</i>	Beulah-London-112079412179026/
<i>Dirk Bikkembergs</i>	dirkbikkembergs
<i>Manolo Blahnik</i>	ManoloBlahnikOfficial
<i>Ozward Boateng</i>	ozwardboatengofficial

<i>BodyMap</i>	bodymapvzw
<i>Boglioli</i>	BoglioliOfficial
<i>Bonia (fashion)</i>	BoniaFashion
<i>Bontoni</i>	Bontoni
<i>Borsalino</i>	Borsalinoworld
<i>Hugo Boss</i>	hugoboss
<i>Bottega Veneta</i>	bottegaveneta
<i>Tony Bowls</i>	tonybowls
<i>BP Studio</i>	BPStudioWorld
<i>Veronique Branquinho</i>	veronique.branquinho.official
<i>Malan Breton</i>	MalanBreton
<i>Thom Browne</i>	THOMBROWNE.newyork
<i>Brunello Cucinelli (company)</i>	BrunelloCucinelli
<i>Bulgari</i>	Bulgari
<i>Burberry</i>	Burberry
<i>Nigel Cabourn</i>	NigelCabourn
<i>Cacharel</i>	Cacharelparis
<i>Calibre (Menswear)</i>	calibreaustralia
<i>Calvin Klein</i>	CalvinKlein
<i>Camilla and Marc</i>	camillaandmarc
<i>Pierre Cardin</i>	PierreCardinTurkiye
<i>Jane Carr (fashion designer)</i>	JANECARR11
<i>Carrera y Carrera</i>	carrera
<i>Cartier (jeweler)</i>	Cartier
<i>Dean and Dan Caten</i>	pages/Dean-and-Dan-Caten/109702452390227
<i>Roberto Cavalli</i>	RobertoCavalli
<i>Cerruti</i>	cerruti.official
<i>Hussein Chalayan</i>	pages/Hussein-Chalayan/10789598271906
<i>Chanel</i>	chanel
<i>Childrensalon</i>	childrensalon
<i>Chloé</i>	chloe.fashion.bags.perfume
<i>Matthew Christopher</i>	matthewbridalcouture
<i>Clements Ribeiro</i>	Clements-Ribeiro-145429422215300/
<i>Colcci</i>	colcciofficial

<i>Comme des Garçons</i>	pages/comme-des-Garçons/105565922809708
<i>Common Projects</i>	commonprojectsofficial
<i>Jasper Conran</i>	jasperconran
<i>Cornelia James</i>	CorneliaJamesLtd
<i>Corneliani</i>	corneliani
<i>CuteCircuit</i>	CuteCircuit
<i>Cutler and Gross</i>	Cutler-and-Gross-10929922121/
<i>Damiani (jewelry company)</i>	damianiofficial
<i>DAS Collection</i>	dascollection
<i>Dege & Skinner</i>	DegeSkinner
<i>Alessandro Dell'Acqua</i>	AllessandroDellAcqua
<i>Ann Demeulemeester</i>	AnnDemeulemeester
<i>Dior Homme</i>	Dior
<i>DKNY</i>	DKNY
<i>Dolce & Gabbana</i>	DolceGabbana
<i>Drake's (haberdashers)</i>	drakeslondon
<i>Duvelleroy</i>	duvelleroy
<i>Equipment (clothing brand)</i>	EquipmentWorld
<i>Ermanno Scervino</i>	ErmannoScervino
<i>Escada</i>	ESCADA
<i>Etro</i>	ETRO
<i>Façonnable</i>	faconnable
<i>Fendi</i>	Fendi
<i>Gianfranco Ferré</i>	pages/Gianfranco-Ferré/111824342168174
<i>Fiorucci</i>	fiorucci.parfums
<i>Tom Ford</i>	tomford
<i>Franck Sorbier</i>	FranckSorbier
<i>Frey Wille</i>	pages/Frey-Wille/136090993081939
<i>Holly Fulton</i>	STUDIOFULTON
<i>Katie Gallagher</i>	officialkatiegallagher
<i>John Galliano</i>	pages/John-Galliano/113498678660336
<i>Jean Paul Gaultier</i>	jean.paul.gaultier
<i>Gieves & Hawkes</i>	Gieveslondon
<i>Giusto Manetti Battiloro</i>	giustomanettibattiloro
<i>Givenchy</i>	Givenchy

<i>Gottex</i>	gottexswimwear
<i>Goyard</i>	Goyard
<i>Gucci</i>	GUCCI
<i>Hackett London</i>	HackettLondon
<i>Kevan Hall</i>	KevanHallSport
<i>Halston</i>	HALSTONHERITAGE
<i>Tim Hamilton (designer)</i>	timhamiltonny
<i>Hardy Amies (fashion house)</i>	hardyamieslndn
<u>Haus Alkire</u>	HAUSALKIRE
<i>Heavy Eco</i>	heavyeco
<i>Helmut Lang (fashion brand)</i>	HELMUTLANG.FB
<i>Alexandre Herchcovitch</i>	pages/Alexandre- Herchcovitch/109522119074435
<i>Hermès</i>	hermes
<i>Carolina Herrera (fashion designer)</i>	CarolinaHerreraNY
<i>Tommy Hilfiger (company)</i>	tommyhilfiger
<i>Sherri Hill</i>	sherrihilldresses
<i>Georges Hobeika</i>	GeorgesHobeikaCouture
<i>House of Flora</i>	houseofflora78
<i>Iceberg (fashion house)</i>	IcebergOfficial
<i>Isabel Marant</i>	isabelmarant
<i>Jil Sander (brand)</i>	pages/Jil-Sander/109715969047306
<i>Jimmy Choo Ltd</i>	JimmyChoo
<i>John Varvatos (company)</i>	JohnVarvatos
<i>Christopher Kane</i>	ChristopherKaneOfficial
<i>Karen Kane</i>	KarenKaneClothing
<i>Kate Spade New York</i>	katespadeny
<i>Mary Katrantzou</i>	Mary-Katrantzou-115328501868015/
<i>Kenzo (brand)</i>	kenzoparis
<i>Khaadi</i>	khaadi
<i>Kiton</i>	Kiton-130087730411396/
<i>L.K.Bennett</i>	LKBennettLondon
<i>Christian Lacroix</i>	christianlacroix
<i>Karl Lagerfeld</i>	karllagerfeld
<i>Lagos (jewelry)</i>	lagosjewelry

<i>Derek Lam</i>	DEREKLAM
<i>Lanvin (company)</i>	LANVINOfficial
<i>Lardini</i>	LardiniOfficial
<i>Guy Laroche</i>	Guy-Laroche-141936919198952/
<i>Larusmiani</i>	Larusmiani
<i>Judith Leiber</i>	JudithLeiber
<i>Léo Marciano</i>	marciano
<i>Monique Lhuillier</i>	OfficialMoniqueLhuillier
<i>Carol Lim</i>	eucarolima
<i>J.Lindeberg</i>	jlindebergofficial
<i>LOEWE (fashion brand)</i>	LOEWE
<i>Longchamp (company)</i>	longchamp
<i>Loro Piana</i>	LoroPiana
<i>Christian Louboutin</i>	christainlouboutin
<i>Luxottica</i>	Luxottica
<i>Bruno Magli</i>	BrunoMagliSpa
<i>Tomas Maier</i>	tomas-maier-134411056587544/
<i>Mainbocher</i>	mainbocher
<i>Maison Margiela</i>	maisonmargiela
<i>Marchesa (brand)</i>	MarchesaFashion
<i>Marimekko</i>	marimekkodesignhouse
<i>Marina Rinaldi</i>	marinarinaldiofficial
<i>Mario Hernández (fashion house)</i>	MARIO-HERNANDEZ-61354397854/
<i>Lana Marks</i>	lanamarkshandbags
<i>Marni (clothing)</i>	marni
<i>La Martina</i>	LaMartinaPolo
<i>Mataano</i>	mataano
<i>Maui Jim</i>	OfficialMauiJim
<i>Max Mara</i>	maxmara
<i>Stella McCartney</i>	stellamcartney
<i>Michael Kors (brand)</i>	michaelkors
<i>Michael Michalsky</i>	pages/Michael- Michalsky/197186457024480
<i>Carlos Miele</i>	CarlosMiele
<i>Missoni</i>	missoni
<i>Miu Miu</i>	MiuMiu

<i>Moschino</i>	pages/Moschino/10938020241057
<i>Mouawad</i>	mouawad
<i>Moynat</i>	MOYNAT
<i>Mulberry (company)</i>	Mulberry
<i>Charlotte Olympia</i>	Charlotte-Olympia-12315443436710/
<i>Oroton</i>	orotonstudio
<i>Rick Owens</i>	rick.owens.corp
<i>Oxford Clothes</i>	OxfordClothes
<i>Cesare Paciotti</i>	CesarePaciottiOfficial
<i>Paule Ka</i>	paulekaparis
<i>Perry Ellis (brand)</i>	PerryEllis
<i>PETROUMAN</i>	PETROUMAN-281405838940/
<i>Philipp Plein</i>	PHILIPP.PLEIN
<i>Poltrona Frau</i>	PoltronaFrau
<i>Pomellato</i>	Pomellato
<i>Henry Poole & Co</i>	HenryPooleCo15SavileRow
<i>Ports 1961</i>	Ports1961
<i>Prada</i>	Prada
<i>Pringle of Scotland</i>	PringleScotland
<i>Emilio Pucci</i>	emiliopucci
<i>Qeelin</i>	QeelinJewellery
<i>Paco Rabanne</i>	pacorabanne
<i>Raketa</i>	raketaklaipeda
<i>Ralph & Russo</i>	ralphandrusso
<i>Ralph Lauren Corporation</i>	RalphLauren
<i>Rayne (shoe company)</i>	
<i>Oscar de la Renta</i>	oscardelarenta
<i>Nina Ricci (brand)</i>	ninaricci
<i>John Richmond (fashion designer)</i>	OfficialJohnRichmond
<i>RM (clothing)</i>	RM.Modaec
<i>Rochas</i>	RochasOfficialPage
<i>Rodarte</i>	OfficialRodarte
<i>Narciso Rodriguez</i>	NarcisoRodriguezStudio
<i>Roland Mouret</i>	rolandmouret
<i>Sergio Rossi</i>	sergiorossi

<i>Christian Roth</i>	ChristianRothEyewear
<i>Rubin & Chapelle</i>	RUBINCHAPELLE
<i>Rue du Mail</i>	ruedumail
<i>Sonia Rykiel</i>	SONIARYKIEL
<i>S.E.H Kelly</i>	sehkelly
<i>Safilo</i>	safilo
<i>Salvatore Ferragamo S.p.A.</i>	SalvatoreFerragamo
<i>Samantha Thavasa</i>	
<i>Sant Blanc</i>	santblanc
<i>Sass & bide</i>	sassandbide
<i>Satya Paul</i>	SatyaPaulIndia
<i>Elsa Schiaparelli</i>	pages/Elsa-Schiaparelli/109584515734307
<i>Sealup</i>	originalsealup
<i>Sermoneta Gloves</i>	sermonetagloves
<i>Shanghai Tang</i>	shangaitang
<i>Shiatzy Chen</i>	shiatzychen.tw
<i>Shreve & Co.</i>	schreveandco
<i>SINOIAN</i>	sinoian
<i>John Smedley (industrialist)</i>	JohnSmedleyKnitwear
<i>Paul Smith (fashion designer)</i>	paulsmithdesign
<i>Smythson</i>	Smythsons
<i>SONGZIO</i>	songziocollection
<i>Luciano Soprani</i>	luciano.soprani.it
<i>Garò Sparo</i>	garosparoaatelier
<i>St. John (clothing)</i>	stjohnsicecaps
<i>Star in the Hood (company)</i>	starinthehood
<i>Strathberry</i>	Strathberry
<i>Jill Stuart</i>	JillStuartIntl
<i>Stylenite</i>	StyleNite
<i>Anna Sui</i>	annasuiimage
<i>Vivienne Tam</i>	viviennetamnyc
<i>TechnoMarine</i>	echnomarine
<i>The Row (fashion label)</i>	therow
<i>Threeasfour</i>	threeASFOUR-78722474228
<i>Tibi (fashion brand)</i>	
<i>Tiffany & Co.</i>	Tiffany

<i>Tori Richard</i>	ToriRichardLtd
<i>Tory Burch (company)</i>	toryburch
<i>Trussardi</i>	trussardi
<i>Turnbull & Asser</i>	Turnbull-Asser-459281314197490/
<i>Anthony Vaccarello</i>	ANTHONY-VACCARELLO- 329000855265/
<i>Valentino SpA</i>	valentino
<i>Kris Van Assche</i>	krisvanasschepage
<i>Dries van Noten</i>	driesvannoten
<i>Alexandre Vauthier</i>	pages/Alexandre- Vauthier/144333142285245
<i>Versace</i>	versace
<i>Louis Vuitton</i>	LouisVuitton
<i>Alexander Wang (designer)</i>	AlexanderWangNY
<i>Vera Wang</i>	VeraWang
<i>Vivienne Westwood</i>	VivienneWestwoodOfficial
<i>Matthew Williamson</i>	WorldofMatthewWilliamson
<i>Wooyoungmi</i>	pariswooyoungmi
<i>Jason Wu</i>	jasonwustudio
<i>Wunderkind (fashion)</i>	wunderkind
<i>Yohji Yamamoto</i>	pages/Yohji-Yamamoto/112278815456032
<i>Yeohlee</i>	YEOHLEE
<i>Yuvna Kim</i>	yuvnakimpage
<i>Giuseppe Zanotti</i>	pages/Giuseppe-Zanotti/253388651406684
<i>Ermenegildo Zegna</i>	ermenegildozegna

Sitografia

- [1] Hootsuite & We Are Social (2016), «Digital 2016 Global Digital Overview,» [Online]. Available: <https://datareportal.com/reports/digital-2016-global-digital-overview>.
- [2] Hootsuite & We Are Social (2021), «Digital 2021 Global Digital Overview,» [Online]. Available: <https://datareportal.com/reports/digital-2021-global-overview-report>.
- [3] A. Langone, Ricerca e analisi di tweet per il mondo della moda, 2016.
- [4] Wikipedia, «Facebook - Wikipedia,» [Online]. Available: <https://it.wikipedia.org/wiki/Facebook>.
- [5] Wikipedia, «Wikipedia - Twitter,» [Online]. Available: <https://it.wikipedia.org/wiki/Twitter>.
- [6] Twitter, «Twitter API,» [Online]. Available: <https://developer.twitter.com/en/docs/twitter-api>.
- [7] Twitter, «Rules and policies API,» [Online]. Available: <https://help.twitter.com/it/rules-and-policies/twitter-api>.
- [8] Twitter, «Twitter rate limit standard API,» [Online]. Available: <https://developer.twitter.com/en/docs/twitter-api/v1/rate-limits>.
- [9] Facebook Inc., «Api Graph,» [Online]. Available: <https://developers.facebook.com/docs/graph-api/overview>.
- [10] Wikipedia, «Rest,» [Online]. Available: https://it.wikipedia.org/wiki/Representational_State_Transfer.
- [11] Facebook, «Facebook Login,» [Online]. Available: <https://developers.facebook.com/docs/facebook-login/overview>.
- [12] Facebook Inc., «Access Tokens,» [Online]. Available: <https://developers.facebook.com/docs/facebook-login/access-tokens/>.
- [13] Facebook Inc., «Permissions Reference,» [Online]. Available: <https://developers.facebook.com/docs/permissions/reference>.

- [14] Facebook Inc., «App review,» [Online]. Available:
<https://developers.facebook.com/docs/app-review/introduction>.
- [15] Facebook Inc., «Server-to-Server App Review,» [Online]. Available:
<https://developers.facebook.com/docs/app-review/resources/sample-submissions/server-to-server/>.
- [16] Facebook Inc., «Page Public Content Access,» [Online]. Available:
https://developers.facebook.com/docs/apps/features-reference/page-public-content-access/?locale=it_IT.
- [17] Facebook Inc., «Enhanced Developer App Review and Graph API 3.0 Now Live,» [Online]. Available: <https://developers.facebook.com/blog/post/2018/05/01/enhanced-developer-app-review-and-graph-api-3.0-now-live/>.
- [18] Facebook Inc., «Verifica dell'azienda,» [Online]. Available:
<https://www.facebook.com/business/help/2058515294227817?id=180505742745347>.
- [19] Facebook Inc., «API Rate Limit,» [Online]. Available:
<https://developers.facebook.com/docs/graph-api/overview/rate-limiting/>.
- [20] R. Yamashita, «Facebook4J,» [Online]. Available: <https://facebook4j.github.io/>.
- [21] Facebook Inc., «Business Management APIs,» [Online]. Available:
<https://developers.facebook.com/docs/business-management-apis>.