Sii-Mobility

Supporto di Interoperabilità Integrato per i Servizi al Cittadino e alla Pubblica Amministrazione

Trasporti e Mobilità Terrestre, SCN_00112

Deliverable ID: DE4.6a
Titolo: Strumenti di Social Intelligence, acquisizione dati e servizi on demand.

Data corrente	10-01-2017
Versione (solo il responsabile puo' cambiare versione	0.3
Stato (draft, final)	Final
Livello di accesso (solo consorzio, pubblico)	Pubblico
WP	WP4
Natura (report, report e software, report e HW..)	Report e software
Data di consegna attesa	M12, Dicembre 2016
Data di consegna effettiva	10-01-2017
Referente primario, coordinatore del documento	UNIFI: DISIT
Contributor	Gianni Pantaleo, gianni.pantaleo@unifi.it Imad Zaza, imad.zaza@unifi.it
Coordinatore responsabile del progetto	Paolo Nesi, UNIFI, paolo.nesi@unifi.it

Sommario

1 Executive Summary

L’uso di strumenti che analizzano i dati provenienti da social network e/o blog è oramai diffuso. Fra le social network più diffuse come Facebook, Twitter, G+, etc., ve ne sono alcune più o meno adatte a poter essere utilizzate per fini di ricerca e di analisi. Fra queste Twitter è una delle più interessanti per le sue caratteristiche di apertura e velocità di reazione della sua utenza. In letteratura, soluzioni basate sull’analisi di dati provenienti da Twitter sono state utilizzate per: il rilevamento dell’arrivo di nuove droghe sul mercato, l’identificazione precoce di eventi e disastri, per la definizione di modelli e soluzioni di capaci di effettuare delle previsioni, per l’analisi dell’apprezzamento di prodotti e persone (in termini di sentiment negativo, positivo, neutro, etc.), per lo studio della risposta ad eventi di intrattenimento televisivo, per la stima delle dimensioni della folla e/o per le predizioni del numero di persone coinvolte in grandi eventi, per la predizione degli andamenti in borsa, ecc. In sostanza, alcuni dati estratti da Twitter, opportunamente elaborati, possono essere sfruttati per calcolare metriche e definire modelli matematici specifici che possono essere utilizzati come strumenti di previsione, diagnosi precoce e per l'analisi della risposta sociale. Ovviamente sono risultati che possono avere un grande valore, oppure un valore limitato dipendentemente dalla correlazione fra la massa delle persone e l’utenza di Twitter.

La maggior parte delle metriche basate su dati provenienti da Twitter si fondano sul conteggio del numero di tweet, del numero di retweet, del numero di follower/amici, il numero di commenti, le relazioni fra utenti, e molti altri parametri che possono essere ottenuti con svariate, e più e o meno complesse elaborazioni. Su questa base, DISIT lab dell’Università degli studi di Firenze ha sviluppato la famiglia di strumenti Twitter Vigilance che oramai sono attivi 24 ore su 24 dall’aprile 2015. In questo periodo sono stati raccolti e analizzati oltre 200 milioni di Tweet per scopi di ricerca. Twitter Vigilance è uno strumento che offre servizi di “intelligence” per la creazione di cruscotti e viste personalizzate per lo studio di eventi e tendenze tramite metriche derivate da Twitter e consente la creazione di nuovi modelli per la previsione, la diagnosi precoce, la valutazione e il monitoraggio, in svariati domini applicativi, definibili dall’utente stesso.

Twitter Vigilance colleziona in modo automatico i dati e su questi effettua operazioni di data mining del contenuto. In accordo alla terminologia di Twitter Vigilance, l’utente può creare dei “Canali” di ascolto, dove ogni Canale di Twitter Vigilance può essere configurato per monitorare un gruppo di chiavi di ricerca su Twitter.com con una sintassi espressiva ed efficace. Dall’interfaccia utente è possibile ottenere direttamente l’andamento di alcune metriche di volume collegate a queste ricerche e molte altre informazioni e andamenti. Alcuni dei Canali che sono attivi su Twitter Vigilance, o che lo sono stati in passato per un certo periodo di tempo, sono accessibili e

sono a disposizione del pubblico tramite la pagina web http://www.disit.org/tv/. E' proprio l'utente che definisce il Canale che può decidere se renderlo accessibile per il pubblico o meno. Inoltre, la pagina di riferimento per informazioni e news su Twitter Vigilance è http://www.disit.org/6693.
Twitter Vigilance viene utilizzato per il monitoraggio dei servizi della città Firenze, per molti aspetti anche a livello regionale, nazionale e/o internazionale; sempre per il controllo della risposta dell'utenza rispetto a eventi critici reali e potenziali, per la valutazione dei servizi di mobilità e di trasporto, per la risposta alle problematiche ambientali e meteo, per la valutazione dei canali e modelli di comunicazione, etc. Twitter Vigilance fornisce una serie di strumenti di analisi e di soluzioni di base ed avanzati per il controllo di metriche basate su dati che provengono da Twitter.
Il deliverable e' stato realizzato in inglese per poter fornire anche a utenti non italiani la possibilita' di comprendere i meccanismi di Twitter Vigilance.

2 Introduction and Objectives

2.1 Acronyms

TV	Twitter Vigilance
OSN	Open Social Network

2.2 Context

The World Wide Web has become an active publishing system and is a rich source of information, thanks to contributions of hundreds of millions of Web users. The growths of online Social Networks in scale and amount of information are immense in recent years. Part of this public expression is carried out on social networking and social sharing sites (Twitter, Facebook, YouTube, etc.), part of it on independent Web sites powered by content management systems (CMSs, including blogs, wikis, news sites with comment systems, Web forums). Content published on this range of Web applications includes information that is newsworthy today or valuable to tomorrow's historians[1].
The analyses of the structure of online social networks have thus drawn much research interests[2]. Before the analyses, the information and the characteristics of the structure have to be obtained. However, the complexity of today's web technologies imposes challenges for collecting the data. The increasing popularity of online social networks (OSNs) has gathered hundreds of millions of users. OSNs have become a platform for people to easily communicate and share information,

particularly with the sophisticate smartphones. Since the structures of OSNs will be able to reflect the real-life society in certain extend, the structure and the information shared in OSNs are of interests for different communities. For instance, sociologists regard OSNs as a venue for collecting relationship data and study online human behaviors. Marketers, in contrast, seek to exploit information about how messages spread so as to design viral marketing strategies. For network engineers, understanding OSNs improves the design of interconnected systems so as to provide better user experience. In order to analyze the structure of an OSN, information regarding the network structure is needed.

The depth and quality of data that can be harvested from Social Media Monitoring tools has evolved significantly in the last years. In [3] has been evaluated the performance of some tools across the following general criteria: Query Set Up, Data Quality, Ease of Data Management, and User Interface & User Experience. Results of that comparison are the following:

- Alterian: Identify common themes among ostensibly disparate conversation
- Brandwatch: Hands on raw data management
- MutualMind: Threaded Facebook discussions
- Radian6: Sophisticated built-in engagement tool; robust dashboard
- Synthesio: Extremely flexible tool all around

What emerged were two distinct categories under which the tools might be classified:

- **Real Time Monitoring and Community Management:** tools that fall under this category seem best suited for monitoring and managing social media communities on a day-to-day basis. This tools are characterized by easily customizable dashboards to make the data a bit more digestible and by the ability to connect various social media accounts.
- **Research & Analysis:** these tools make the analysis of vast quantities of conversation data more manageable, and are generally better suited to monitoring social media performance over the long term. Access to historical data also makes these tools ideal for the development and evaluation of social media strategies.

Social media is especially important for research into computational social science that investigates questions using quantitative techniques (e.g., computational statistics, machine learning and complexity) and so-called big data for data mining and simulation modeling[4]. Social media scraping and analytics provides a rich source of academic research challenges for social scientists, computer scientists and funding bodies. Challenges include:

- **Scraping**—although social media data is accessible through APIs, due to the commercial value of the data, most of the major sources such as Facebook and Google are making it increasingly difficult for academics to obtain comprehensive access to their 'raw' data; very few social data sources provide affordable data offerings to academia and researchers.
- **Data cleansing**—cleaning unstructured textual data (e.g., normalizing text), especially high-frequency streamed real-time data, still presents numerous problems and research challenges.
- **Holistic data sources**—researchers are increasingly bringing together and combining novel data sources: social media data, real-time market & customer data and geospatial data for analysis.
- **Data protection**—once you have created a 'big data' resource, the data needs to be secured, ownership and IP issues resolved (i.e., storing scraped data is against most of the publishers' terms of service), and users provided with different levels of access; otherwise, users may attempt to 'suck' all the valuable data from the database.
- **Data analytics**—sophisticated analysis of social media data for opinion mining (e.g., sentiment analysis) still raises a myriad of challenges due to foreign languages, foreign words, slang, spelling errors and the natural evolving of language.
- **Analytics dashboards**—many social media platforms require users to write APIs to access feeds or program analytics models in a programming language, such as Java. While reasonable for computer scientists, these skills are typically beyond most (social science) researchers. Non-programming interfaces are required for giving what might be referred to as 'deep' access to 'raw' data, for example, configuring APIs, merging social media feeds, combining holistic sources and developing analytical models.
- **Data visualization**—visual representation of data whereby information that has been abstracted in some schematic form with the goal of communicating information clearly and effectively through graphical means. Given the magnitude of the data involved, visualization is becoming increasingly important.

2.3 State of the art of tool for twitter data analysis

The analysis of social networks has had a significant development, especially for commercial applications in marketing and advertising campaigns. In fact, there are many tools that offer paid services for semantic and sentiment analysis to companies that want to evaluate the perception of its brand by users, the trend of a particular advertising campaign and all that can be extracted by the enormous mass data that social networks provide. The analysis tools can be divided into two categories: those that offer only the technology to retrieve information leaving the user the task of analyzing and those that offer a complete service.

2.3.1 Retrieve Only

2.3.1.1 80legs

In the first group we can insert **80legs** [5] that provides a web crawling directly from the website or through the API. 80legs gives the ability to customize crawl by providing a set of options that specific how crawl will run. These options are:

1. A list of URLs which tell 80legs where to start the crawl
2. A list of criteria to explain to 80legs what data to scrape from each crawled URL, as well as what URLs to crawl next
3. Other options to control the crawl, such as the number of total URLs to crawl.

2.3.1.2 Visual Web Riper

Another tool that allows the retrieval of information from the web is **Visual Web Ripper** [6] that contains a wealth of advanced features that enables user to harvest data from even the most difficult websites. Visual Web Ripper can be configured to download complete content structures, such as product catalogs.

2.3.1.3 Helium Scraper

HeliumScraper [7] can also be inserted in the first group because is a web scraping tool that can be trained to extract specific information from web sites. The results can be exported in a variety of formats.

2.3.1.4 PromptCloud

PromptCloud [8] is a classic web crawling model where is taken the list of sites that user would like crawled and do vertical-specific crawls. It's possible to provide PromptCloud with a list of keywords that gets fed into crawler. The crawler then continuously looks for matching tweets to that list of keywords as tweets gets published. All these tweets are later converted into a structured format with other associated information. In paper [9] authors propose one such tool called **Intention Insider** which has been developed at HP Labs in close collaboration with business units and a few selected customers. The tool can ingest content from online forums or from uploaded files and quickly sift through very large amounts of comments to extract intention information. This information is loaded into a data warehouse to be correlated with other structured data and queried to produce interactive reports and dynamic visualizations that facilitate its exploration at detailed and aggregate levels.

2.3.2 Retrieve and analyze

The second group of tools is much wider because the market requires a comprehensive toolset that directly provide the test results.

2.3.2.1 Openamplify

These include **Openamplify** [10] that is an Natural Language Processing (NLP) analytic engine that processes text to extract valuable knowledge from social media conversations. This tool gives a picture of brand's health or campaign's performance. OpenAmplify analyzes any provided text, structured or unstructured, without the need for training or special vocabularies, returns a set of "signals", each of which describes a specific aspects of the text's meaning, sentiment, intent, style, or other characteristics and delivers the signals ranked and organized in useful ways.

2.3.2.2 Clarabridge Intellgence Platfrom

Clarabridge Intelligence Platform [11] gets a complete view of customers' experience. The Clarabridge Intelligence Platform harnesses all available sources of consumer feedback, including multiple survey types, contact center agent notes, social media, chat, voice, email, warranty notes, and much more. The Clarabridge Intelligence Platform's core functionality includes text analytics, context-sensitive sentiment analysis, linguistic categorization, and emotion detection. Clarabridge Social seamlessly gathers and accurately analyzes any online customer data, whether structured or unstructured, from any social media source and any online review site. Clarabridge Social connects to all popular sites including Facebook, Twitter, Trip Advisor, and Bookings.com, and integrates with social media management software such as Sysomos, Radian6, and many more for a comprehensive view.

2.3.2.3 Bradwatch Analytics

Brandwatch Analytics [12] is a web-based social media monitoring platform designed to allow users to get the most out of the social media data important to their business. It is focused on Customer Experience. Its main use cases are: Brand/reputation management, Finding influences/advocates, Market research, Campaign, Crisis management, Community management, PR, Customer Services, SEO and Lead generation. Channels feature allows to track any public Facebook page or Twitter account without the need for admin rights.

2.3.2.4 Opinion crawl

Opinion Crawl [13] allows Companies and agencies to order in-depth reports monitoring online image of an entity - a company, a product, or an individual. The crawlers process large amounts of various Web sources - blogs, news sites, forums. The reports are produced on an ongoing basis and emailed to the client. The reports are broken by day/week/month, and contain current and trend charts, key concepts associated with the topic, and references to source documents. Sentiment API allows client applications to assess sentiment on a Web page or a piece of text, e.g. a blog comment.

2.3.2.5 Social report

Social Report [14] is a social network analytics solution that allows to track social network accounts just the same way it's possible to track the performance of websites. Social Report tracks and monitors social network accounts and gives user tools to manage marketing initiatives. Social Report offers powerful insights into social accounts: membership trends, activity and engagement, thoughts and feelings of members, their interests, their geographical distribution, education levels, gender, employment, and countless other metrics.

2.3.2.6 Mozenda

Mozenda [15] is a web scraping service used by many well-known brands. The Agent Builder supports the creation of agents that collect specific information from web sites. These are created in a Windows environment and submitted to the service where they are executed. The Web Console allows agents to be run and scheduled and export and publish the results of a search.

2.3.2.7 Beevolve

Beevolve [16] monitors brand mentions, schedules and launch social media posts and measure resulting sales and engagement for those posts.

2.3.2.8 Meltwater

Meltwater's [17] online intelligence platform analyzes digital documents daily to extract precise, timely business insights that help executives understand their markets, engage their customers, and master the new social business environment. Meltwater PR solutions help public relations and marketing communications professionals build brands and drive growth by effectively engaging media influencers. Meltwater social media marketing solutions combine deep social media monitoring with efficient social engagement to help creating more effective marketing campaigns across large social communities.

2.3.2.9 Viralheat

Viralheat's [18] sentiment analysis tool allows user to understand the sentiment of online mentions for business, brands, and products. Identifies the sentiment of social mentions across multiple social platforms and pulled detailed analysis of what users are saying about a product or service. This tool allows view sentiment of social posts in real-time and pull sentiment analytics from Facebook, Twitter, Instagram, and Tumblr.

2.3.2.10 SAS Sentiment Analysis

SAS Sentiment Analysis [19] automatically rates and classifies opinions expressed in electronic text. It collects text inputs from websites, social media outlets and internal file systems, and then puts them in a unified format to assess relevance to predefined topics. Reports identify trends or emotional changes, and an interactive workbench allows subject-matter experts to refine sentiment models.

2.3.2.11 Dataminr

Dataminr [20] transforms real-time data from Twitter and other public sources into actionable signals, identifying the most relevant information in real-time for clients in Finance, the Public Sector, News, Security and Crisis Management. In partnership with Twitter, Dataminr developed and launched Dataminr for News, which alerts journalists to breaking news in advance of traditional sources and is now used by hundreds of news organizations globally. Most recently, Dataminr launched a product for security and crisis management watch centers that warns the world's largest corporations to emerging threats and crises, ensuring that a corporation's physical assets and employees are protected.

2.3.2.12 Tracx

Tracx [21] is a company with a SaaS platform for sophisticated brand marketers who want to do more than monitor their social media presence, but actually manage it. The company provides an end-to-end solution that indexes the entire social web and delivers the most relevant, high impact audiences and conversations by capturing a 360 degree view of activity around a brand.

2.3.2.13 Roialty

ROIALTY [22] is the digital loyalty platform (web, social, mobile) that allows a brand to develop the potential of 'engagement' in social media & digital communities by increasing their involvement through Gamification & rewarding. It gives a boost to awareness, conversions and purchases on the e-commerce and retail channels monitoring the full range of interaction metrics needed to measure the ROI (Return On Investment) of each digital campaign. ROIALTY rewards authenticated users connecting their social profiles and offers them some targeted 'missions', based on their socio-demographics, preferences and interests. Each mission engages the user in content creation & publishing on blogs and social media or in promotion of product/service initiatives through likes and sharing as well as participating in surveys.

2.4 Twitter

Twitter is a social network that deals with free microblogging, devised by the american Jack Dorsey and developed by Obvious Corporation in San Francisco [23]. The service offered to members is the inclusion of messages, called 'tweets', consisting of a maximum of 140 characters.

Since its creation and networking in March 2006, Twitter has taken a prominent role within the set of social networks on the Web, reaching more than 250 million active users [24], that found in the service a quick way to interact with the rest of the world.

In addition to simple text, within the tweet you can enter keywords, called hashtags (preceded by a "#"), and links to other sites, usually abbreviated URL services via shorting. Messages posted by members are by default rendered visible to anyone, whether registered or not in the service, while you can make your tweets private so that they can only be read by authorized persons [25]. The inclusion of the messages is made possible not only through the social network site, but also by a number of external applications and, limited to a few countries, via SMS. The ability to send messages via different devices and applications, is one of the strengths of the social network. Subscribers to the service have the opportunity to follow other registered users: they assume in this case the name of 'followers' and have the ability to view in their own "home page" messages posted by those users. Is also possible to follow lists of users, in other words lists created by other subscribers in which is included a variable number of users. Another important factor for the service is the ability to respond to messages from any other registered user, thus creating conversations online. A member can forward the message to another user who is following, so that it is visible to all their followers. This is called retweet and is reported in messages prefixing characters RT to the original text.

It is should be noted that the conversations and personal status, calculated on a sample of tweets harvested from social networks, almost reach 90% of the total posts, while 37.55% of the total is made up of messages in response to other tweets. As for spam messages and self-promotion (ie tweet purposes only advertising placed by companies) they are limited to 9.6%. From these statistics it is possible to deduce how Twitter has become one of the most effective means to share experiences and how users can use it for communicating with each other, getting closer to the original idea of Jack Dorsey to create a service similar to SMS, but applied to groups of people and available on the web.

Since the beginning, the study of Twitter has proved of huge interest, being a social network popular, dynamic and where users, through their tweets, help to keep the public informed of what is happening around them. In 140 characters of a message are told life experiences both personal and about the world that surrounds users, while the ability to include links to other sites, as well as images and videos, are additional methods to disseminate what is most important in web.

3 Twitter Vigilance

Twitter Vigilance is designed as a platform for the search of messages on Twitter and for the analysis of such messages both from a numerical point of view (to highlight the daily peaks) and from a semantic point of view (for identify what refers peaks of tweets).

The idea of creating a platform for the monitoring of social networks is created within a collaboration between UNIFI DISIT lab, LAMMA and CNR IBINET. The main purpose is to investigate and to build specific and reliable metrics dashboard to monitor weather related Twitters. Since this study was the project of Twitter Vigilance Platform to provide a tool to study the content of the social network that was shared and adapted to different contexts, such as for monitoring city services, critical events and conditions, user behavior, city response to events, etc.. Indeed, the tool that will be made will also be used in other projects related to smart cities.

Specifically, the main objective of the collaboration mentioned above is to analyze whether can be used the information flow of Twitter as a good social indicator of certain weather events, such as severe weather alerts or heat waves. Therefore, in the design of Twitter Vigilance Platform has been introduced the concept of "thematic channel" in which research is collected concerning a certain topic, as can be weather alert or particular events as Expo.

A tool of this kind cannot be separated from the management of users and the distinction of roles: basically will be split functionality between the User and the Administrator. A user must be able to create its channels and view only his own, but may be able to include in its channels every search included in the system even if inserted by another user. The administrator must be able to manage all channels of all users as well as monitor all system activity.

Since the platform is designed to be an analytics tool, a key part of the system is displaying the results of statistics and analyzes performed on messages within the channels. To achieve this purpose it was decided to create a dashboard that can show to the user, through graphs and charts, information that may be useful to explore its analysis.

Gli strumenti della soluzione integrata Twitter Vigilance sono accessibili via WEB e sono adatti per lo studio, la ricerca ed il monitoraggio di social media via Twitter. In particolare, i loro punti accesso sono:

- **Twitter Vigilance main** tool: http://www.disit.org/tv/
- **Real Time Twitter Vigilance**: http://www.disit.org/rttv/

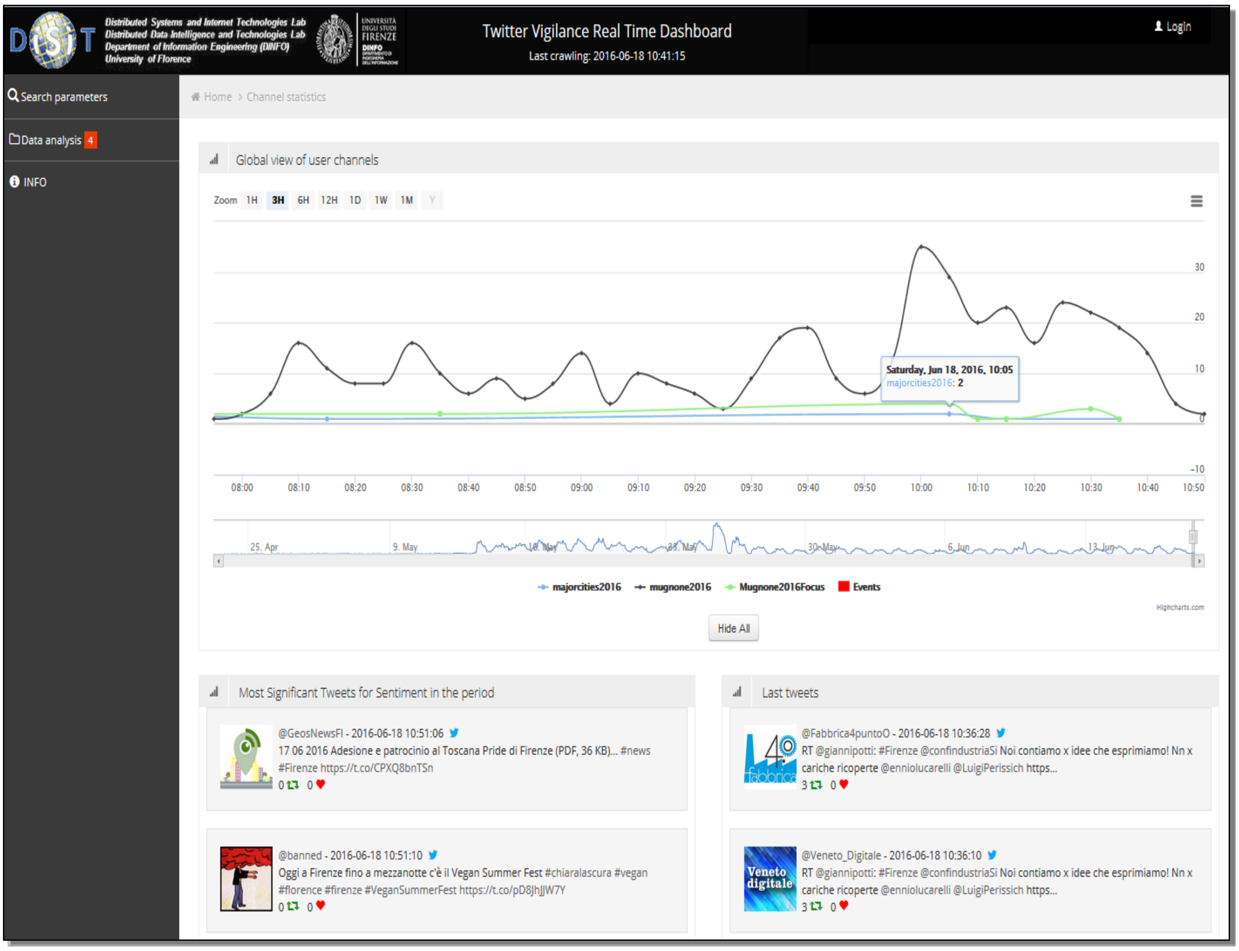
- **Twitter Vigilance Advanced Search** facility based on SOLR: http://tvsolr.disit.org/search/?collection=1

Twitter Vigilance main tool permette di :

(i) gestire, attivare e conFigurere Canali e ricerche;
(ii) seguire le principali tendenze di Twitter tramite i propri canali,
(iii) analizzare i trend delle metriche computate in automatico,
(iv) eseguire analisi sugli utenti e relazioni, e su richiesta
(v) eseguire elaborazioni di natural language processing, NLP, e sentiment analysis, SA, ecc.

La maggior parte di queste valutazioni sono calcolate/aggiornate con cadenza oraria e/o giornaliera, permettendo in questo modo di effettuare previsioni e valutazioni su eventi/accadimenti che hanno dinamiche lente o che sono avvenuti nel passato. In particolare, l'analisi relativa all'NLP e alla SA integra in maniera innovativa tecniche allo stato dell'arte come il parsing sintattico automatico del testo, soluzioni di Part-Of-Specch (POS) tagging, nonché algoritmi di Named Entity Recognition e disambiguazione, insieme a risorse esterne semantiche annotate per la determinazione della polarità di Sentiment. Da questo strumento si possono scaricare informazioni e dati che possono essere rielaborati per creare modelli statistici predittivi.

Quando gli eventi variano velocemente nel tempo, come per esempio per la diagnosi precoce di condizioni critiche, per seguire l'evoluzione di una trasmissione televisiva, per seguire la risposta dell'uditorio rispetto a un dibattito; è necessario attivare delle elaborazioni mirate che seguono il flusso dati in tempo reale. Questo tipo di elaborazione è possibile tramite **Real Time Twitter Vigilance (**http://www.disit.org/rttv/**)** in cui ogni Canale, acquisisce i dati da Twitter e gli elabora in tempo reale effettuando valutazioni statistiche, NLP e SA; fornendo in questo modo direttamente i risultati in forma di grafico e di lista dei Tweet su base temporale e della risposta emotiva agli stimoli ed accadimenti.



Distributed Systems and Internet Technologies Lab
Distributed Data Intelligence and Technologies Lab
Department of Information Engineering (DINFO)
University of Florence

Twitter Vigilance Real Time Dashboard
Last crawling: 2016-06-18 10:41:15

Most Significant Tweets for Sentiment in the period

@GeosNewsFI - 2016-06-18 10:51:06
17 06 2016 Adesione e patrocinio al Toscana Pride di Firenze (PDF, 36 KB)... #news #Firenze https://t.co/CPXQ8bnTSn
0 0

@banned - 2016-06-18 10:51:10
Oggi a Firenze fino a mezzanotte c'è il Vegan Summer Fest #chiaralascura #vegan #florence #firenze #VeganSummerFest https://t.co/pD8JhJJW7Y
0 0

Last tweets

@Fabbrica4puntoO - 2016-06-18 10:36:28
RT @giannipotti: #Firenze @confindustriaSi Noi contiamo x idee che esprimiamo! Nn x cariche ricoperte @enniolucarelli @LuigiPerissich https...
3 0

@Veneto_Digitale - 2016-06-18 10:36:10
RT @giannipotti: #Firenze @confindustriaSi Noi contiamo x idee che esprimiamo! Nn x cariche ricoperte @enniolucarelli @LuigiPerissich https...
3 0

Figure 1: Real Time Twitter Vigilance

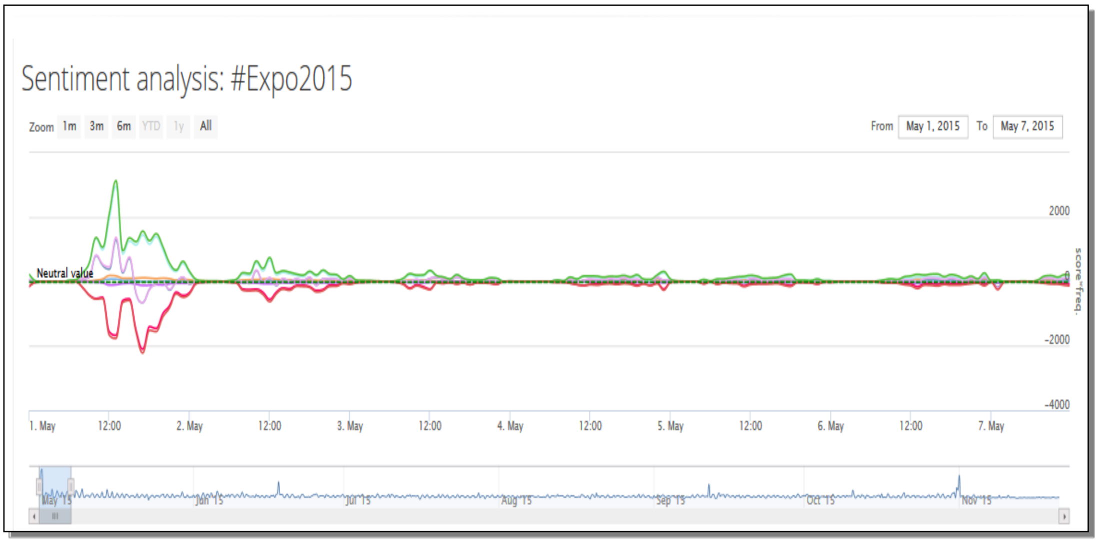
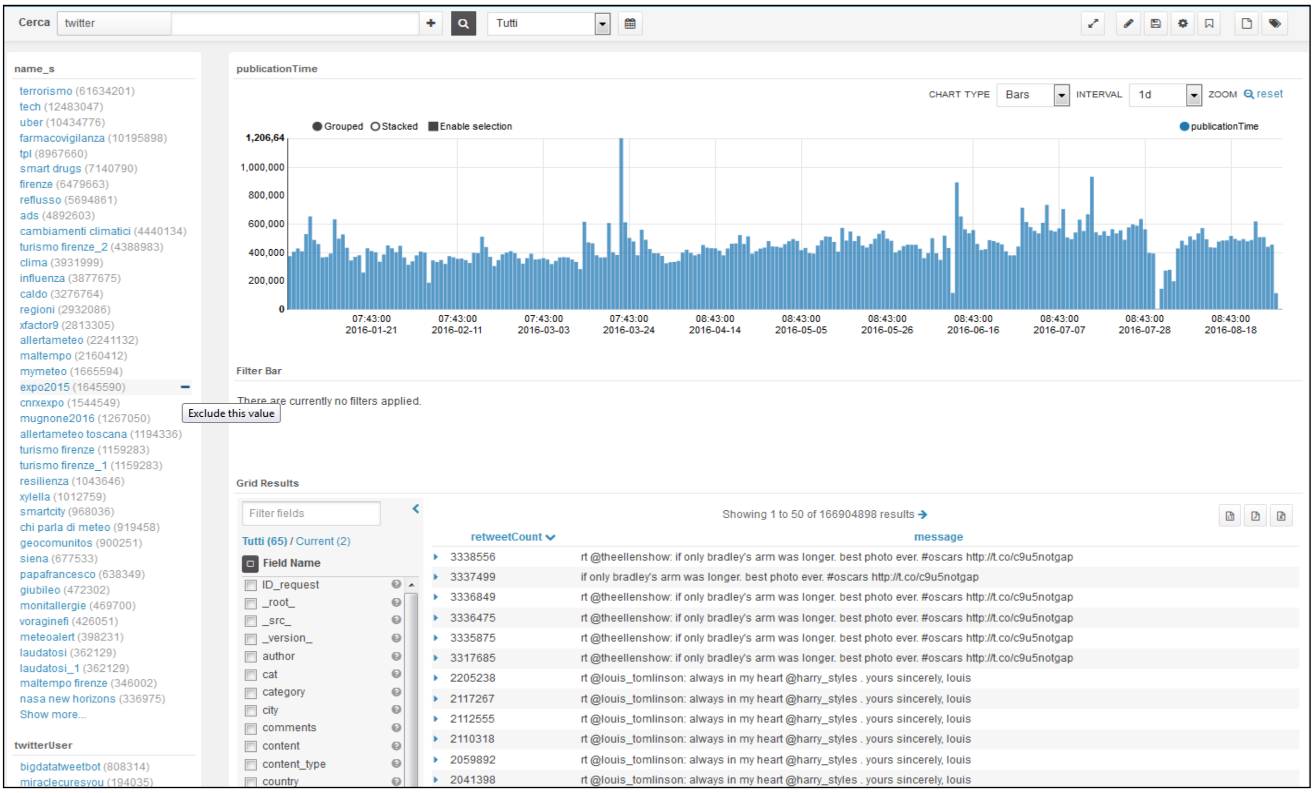


Figure 2: una vista sulla Sentiment Analysis

In alcuni casi, gli analisti di dati social media hanno la necessità di ricercare all'interno dei tweet e dati collezionati combinazioni particolari incrociando canali, ricerche, lingue, citazioni, utenti, periodi temporali, luoghi, etc., effettuando varie combinazioni AND/OR e/o faceted/sfaccettate. A questo fine è stato sviluppato lo strumento **Twitter Vigilance Advanced Search** che permette di navigare nel pool di tutti i tweet collezionati con un indice SOLR, per esempio, http://tvsolr.disit.org/search/?collection=1. Anche in questo caso si possono conFigurere e concordare la creazione di viste specifiche, o anche l'ampliamento del modello di ricerca attuale.



name_s

- terrorismo (61634201)
- tech (12483047)
- uber (10434776)
- farmacovigilanza (10195898)
- tpl (8967660)
- smart drugs (7140790)
- firenze (6479663)
- reflusso (5694861)
- ads (4892603)
- cambiamenti climatici (4440134)
- turismo firenze_2 (4388983)
- clima (3931999)
- influenza (3877675)
- caldo (3276764)
- regioni (2932086)
- xfactor9 (2813305)
- allertameteo (2241132)
- maltempo (2160412)
- mymeteo (1665594)
- expo2015 (1645590)
- cnrxexpo (1544549)
- mugnone2016 (1267050)
- allertameteo toscana (1194336)
- turismo firenze (1159283)
- turismo firenze_1 (1159283)
- resilienza (1043646)
- xylella (1012759)
- smartcity (968036)
- chi parla di meteo (919458)
- geocomunitos (900251)
- siena (677533)
- papafrancesco (638349)
- giubileo (472302)
- monitallergie (469700)
- voraginefi (426051)
- meteoalert (398231)
- laudatosi (362129)
- laudatosi_1 (362129)
- maltempo firenze (346002)
- nasa new horizons (336975)
- Show more...

twitterUser

- bigdatatweetbot (808314)
- miraclecuresyou (194035)

Filter Bar

There are currently no filters applied.

Grid Results

Tutti (65) / Current (2)

- □ ID_request
- □ _root_
- □ _src_
- □ _version_
- □ author
- □ cat
- □ category
- □ city
- □ comments
- □ content
- □ content_type
- □ country

Showing 1 to 50 of 166904898 results

retweetCount	message
3338556	rt @theellenshow: if only bradley's arm was longer. best photo ever. #oscars http://t.co/c9u5notgap
3337499	if only bradley's arm was longer. best photo ever. #oscars http://t.co/c9u5notgap
3336849	rt @theellenshow: if only bradley's arm was longer. best photo ever. #oscars http://t.co/c9u5notgap
3336475	rt @theellenshow: if only bradley's arm was longer. best photo ever. #oscars http://t.co/c9u5notgap
3335875	rt @theellenshow: if only bradley's arm was longer. best photo ever. #oscars http://t.co/c9u5notgap
3317685	rt @theellenshow: if only bradley's arm was longer. best photo ever. #oscars http://t.co/c9u5notgap
2205238	rt @louis_tomlinson: always in my heart @harry_styles . yours sincerely, louis
2117267	rt @louis_tomlinson: always in my heart @harry_styles . yours sincerely, louis
2112555	rt @louis_tomlinson: always in my heart @harry_styles . yours sincerely, louis
2110318	rt @louis_tomlinson: always in my heart @harry_styles . yours sincerely, louis
2059892	rt @louis_tomlinson: always in my heart @harry_styles . yours sincerely, louis
2041398	rt @louis_tomlinson: always in my heart @harry_styles . yours sincerely, louis

Figure 3: Twitter Vigilance Advance search

3.1.1 Frontend Architecture

The front end is designed to be used either by the User that the Administrator: to accomplish that are designed two different architectures for the two types of users. Moreover, has been included a display of some parts of the system also to unauthenticated users, which is simply a derivation of the display designed for authenticated users. As for the user side, the front end must allow to manage the channels associated to the User and see some statistics about the messages downloaded primarily in graphical form. As for the part of Administrator, the frontend must allow the management and viewing statistics for all channels and all searches included in the system, checking the status of background processes and viewing statistics on background processes (in particularly for the process that queries on Twitter).

3.1.1.1 User Frontend

When you enter the front end by unauthenticated or authenticated user the first page that appears is the page "Channel Statistics": This page shows the list of "public" channels both in the form of a graph and in table form (Figure 4). In the column Detail there are two buttons: the first on the left makes access to the detail page of the selected channel, the second show the NLP or sentiment analysis of tweets of the channel if previously triggered.

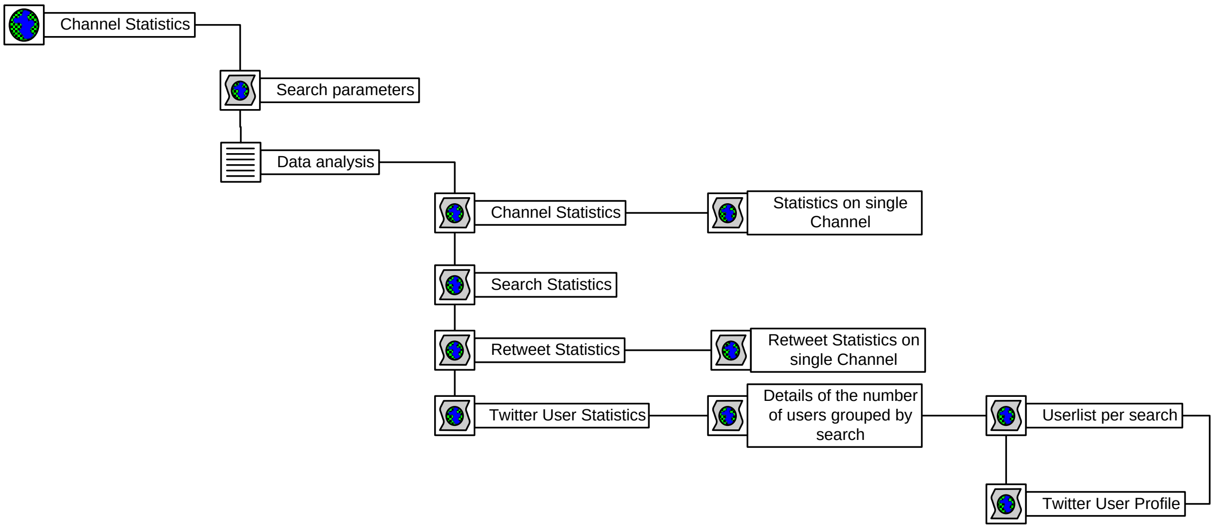
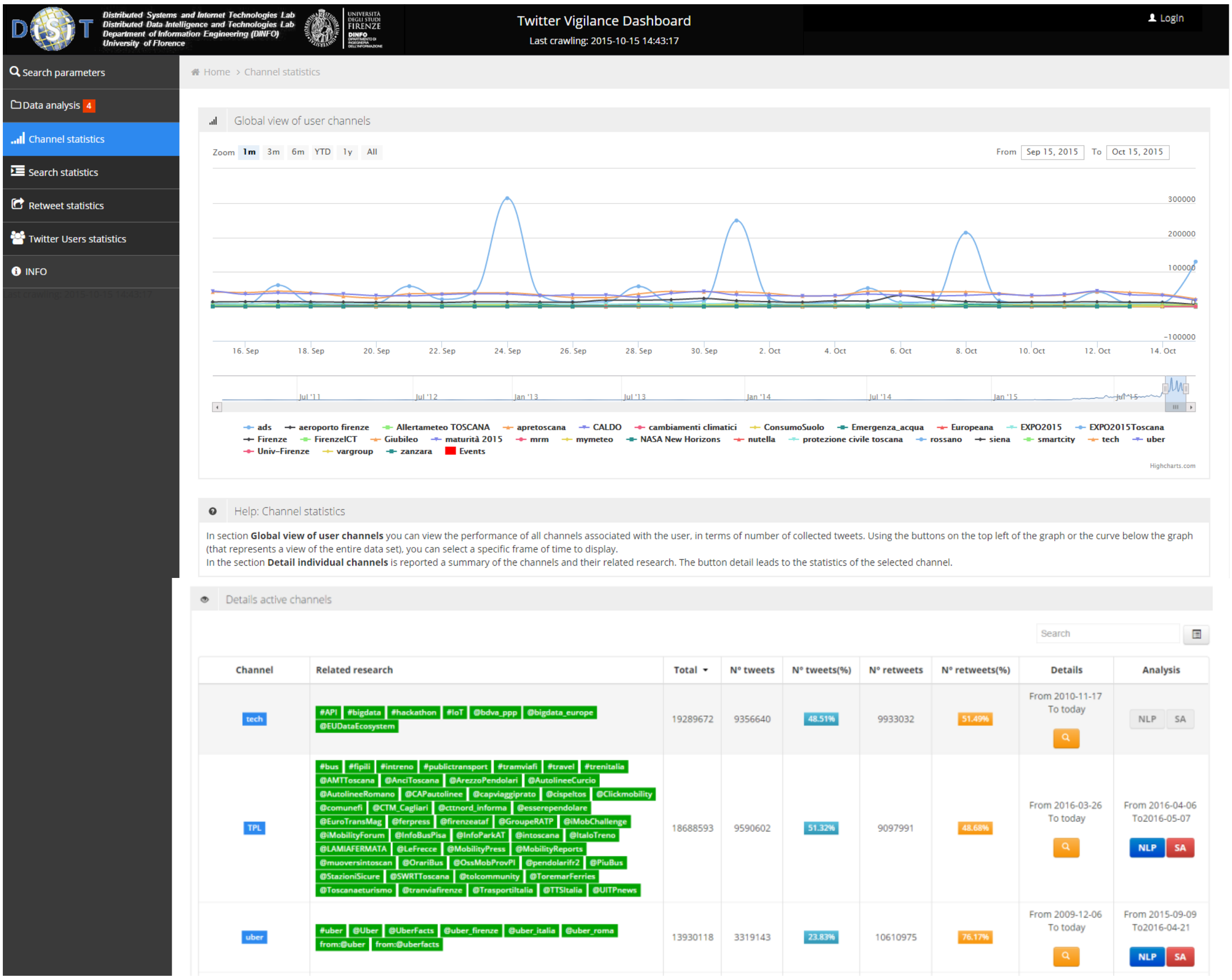


Figure 4: Structure of User Frontend

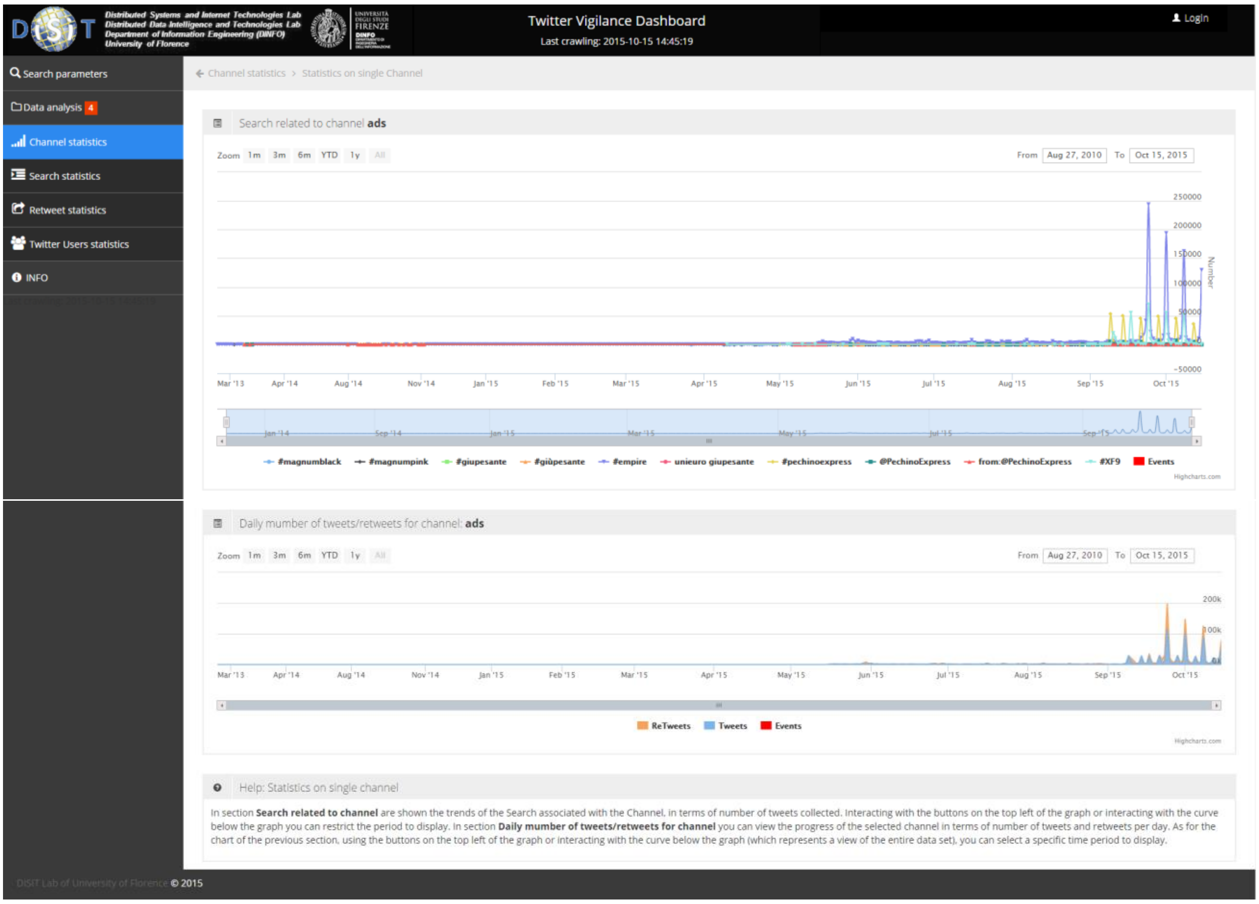


In section **Global view of user channels** you can view the performance of all channels associated with the user, in terms of number of collected tweets. Using the buttons on the top left of the graph or the curve below the graph (that represents a view of the entire data set), you can select a specific frame of time to display.
In the section **Detail individual channels** is reported a summary of the channels and their related research. The button detail leads to the statistics of the selected channel.

Channel	Related research	Total	N° tweets	N° tweets(%)	N° retweets	N° retweets(%)	Details	Analysis
tech	#API #bigdata #hackathon #IoT @bdva_ppp @bigdata_europe @EUDataEcosystem	19289672	9356640	48.51%	9933032	51.49%	From 2010-11-17 To today	NLP SA
TPL	#bus #fipili #intreno #publictransport #tramviafi #travel #trenitalia @AMTToscana @AnciToscana @ArezzoPendolari @AutolineeCurcio @AutolineeRomano @CAPautolinee @capviaggiprato @cispeltos @Clickmobility @comunefi @CTM_Cagliari @cttnord_informa @esserependolare @EuroTransMag @ferpress @firenzeataf @GroupeRATP @iMobChallenge @iMobilityForum @InfoBusPisa @InfoParkAT @intoscana @ItaloTreno @LAMIAFERMATA @LeFrecce @MobilityPress @MobilityReports @muoversintoscan @OrariBus @OssMobProvPI @pendolarifr2 @PiuBus @StazioniSicure @SWRTToscana @tolcommunity @ToremarFerries @Toscanaeturismo @tranviafirenze @TrasportiItalia @TTSItalia @UITPnews	18688593	9590602	51.32%	9097991	48.68%	From 2016-03-26 To today	From 2016-04-06 To2016-05-07 NLP SA
uber	#uber @Uber @UberFacts @uber_firenze @uber_italia @uber_roma from:@uber from:@uberfacts	13930118	3319143	23.83%	10610975	76.17%	From 2009-12-06 To today	From 2015-09-09 To2016-04-21 NLP SA

Figure 5: Channel statistics page

The graph shown in Figure 5 shows the number of tweets per day for each channel associated with the user.



In section **Search related to channel** are shown the trends of the Search associated with the Channel, in terms of number of tweets collected. Interacting with the buttons on the top left of the graph or interacting with the curve below the graph you can restrict the period to display. In section **Daily number of tweets/retweets for channel** you can view the progress of the selected channel in terms of number of tweets and retweets per day. As for the chart of the previous section, using the buttons on the top left of the graph or interacting with the curve below the graph (which represents a view of the entire data set), you can select a specific time period to display.

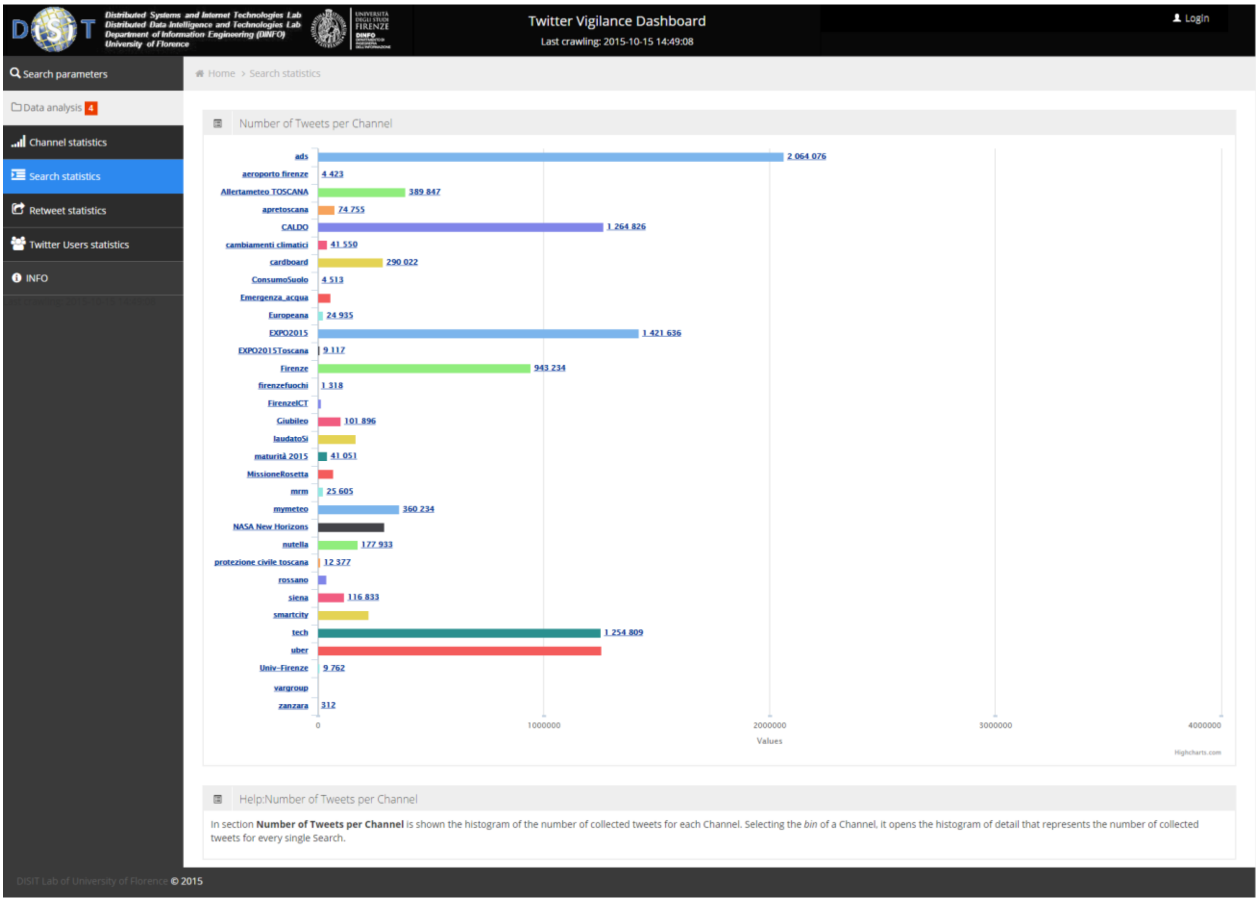
Figure 6: Statistics on single Channel page

Details page of the channel, Figure 6 shows two graphs: the top one shows the number of tweets per day for each search associated with the channel, the second shows the number of tweets and retweets per day for the entire channel.

For the display of this graph was prepared the table "chart_twitter" in which are stored the necessary data by a separate process which will be described later. This table has updated at regular intervals by adding only the values for the downloaded messages since the last update.

3.1.1.2 Search Statistics

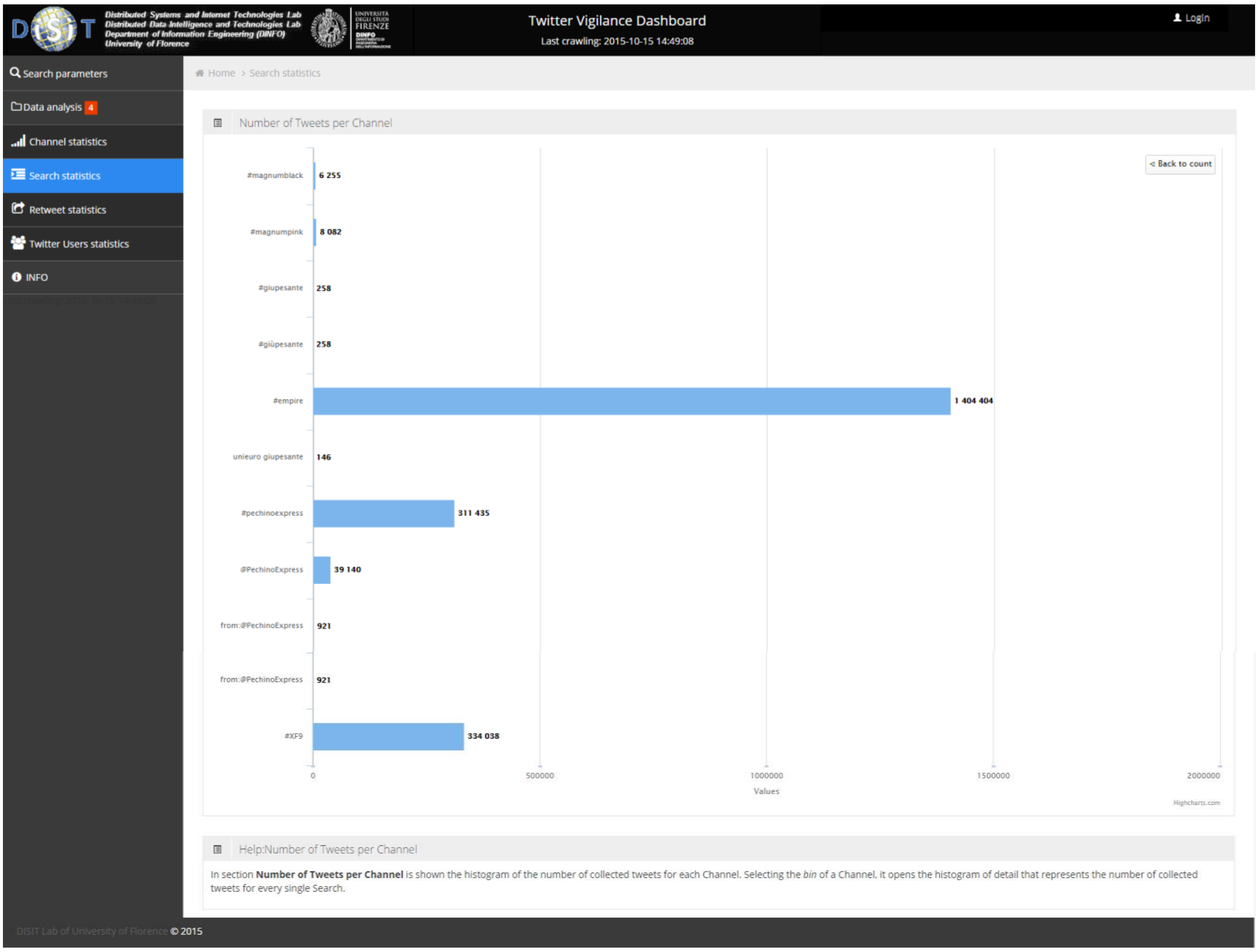
On this page (Figure 7) you can see a histogram that shows the total number of messages for each channel.



In section **Number of Tweets per Channel** is shown the histogram of the number of collected tweets for each Channel. Selecting the *bin* of a Channel, it opens the histogram of detail that represents the number of collected tweets for every single Search.

Figure 7: Search statistics page

Clicking on a bin it is possible display another histogram that shows the total values of the messages of the individual searches that are part of channel, as shown in Figure 8.

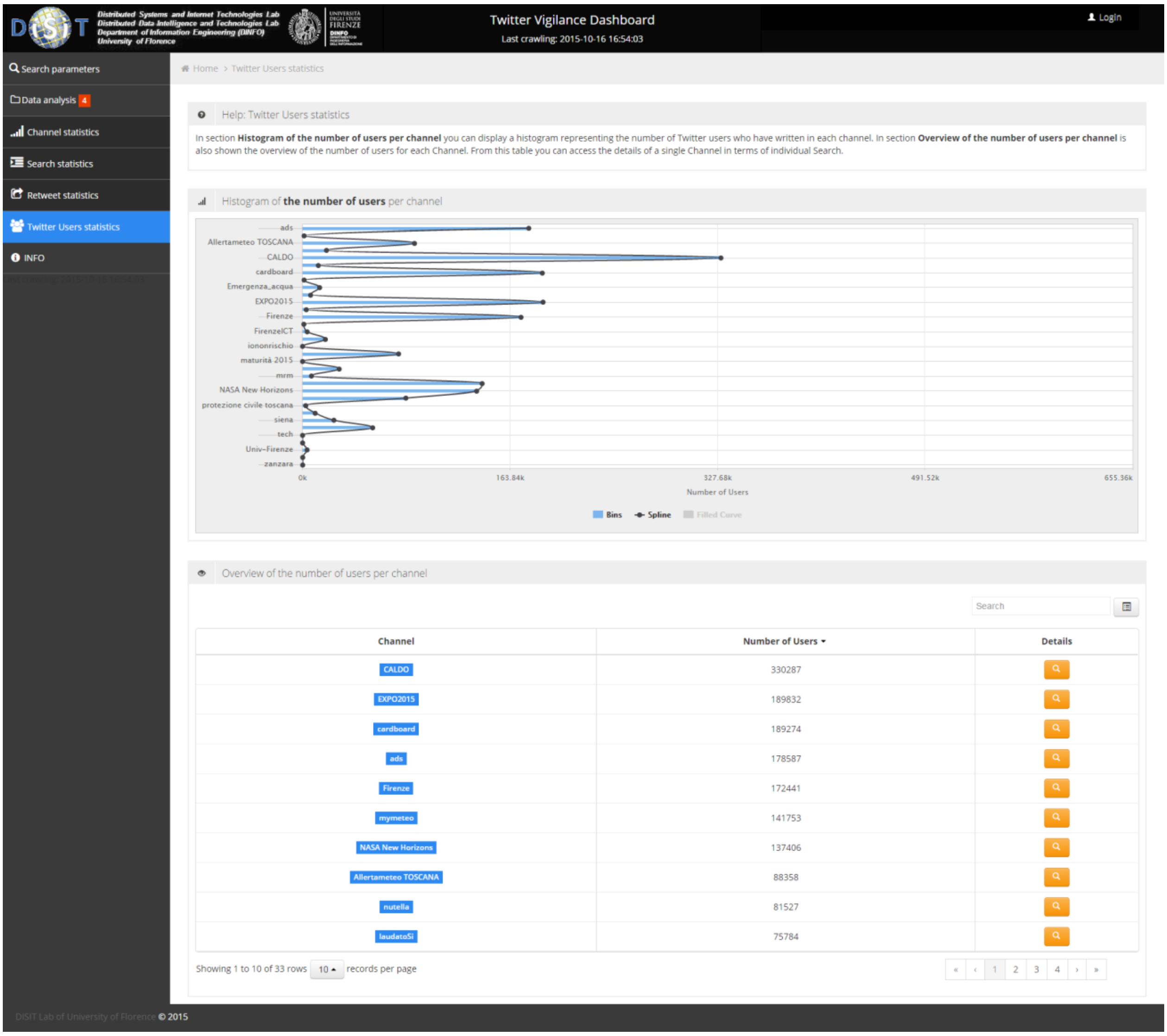


In section **Number of Tweets per Channel** is shown the histogram of the number of collected tweets for each Channel. Selecting the *bin* of a Channel, it opens the histogram of detail that represents the number of collected tweets for every single Search.

Figure 8: Search statistics page: single channel details

3.1.1.3 Twitter User Statistics

In this page (Figure 9) there are a histogram and a table representing the same data, ie the total number of distinct users for each channel.



Twitter Vigilance Dashboard

Last crawling: 2015-10-16 16:54:03

Help: Twitter Users statistics

In section **Histogram of the number of users per channel** you can display a histogram representing the number of Twitter users who have written in each channel. In section **Overview of the number of users per channel** is also shown the overview of the number of users for each Channel. From this table you can access the details of a single Channel in terms of individual Search.

Histogram of **the number of users** per channel

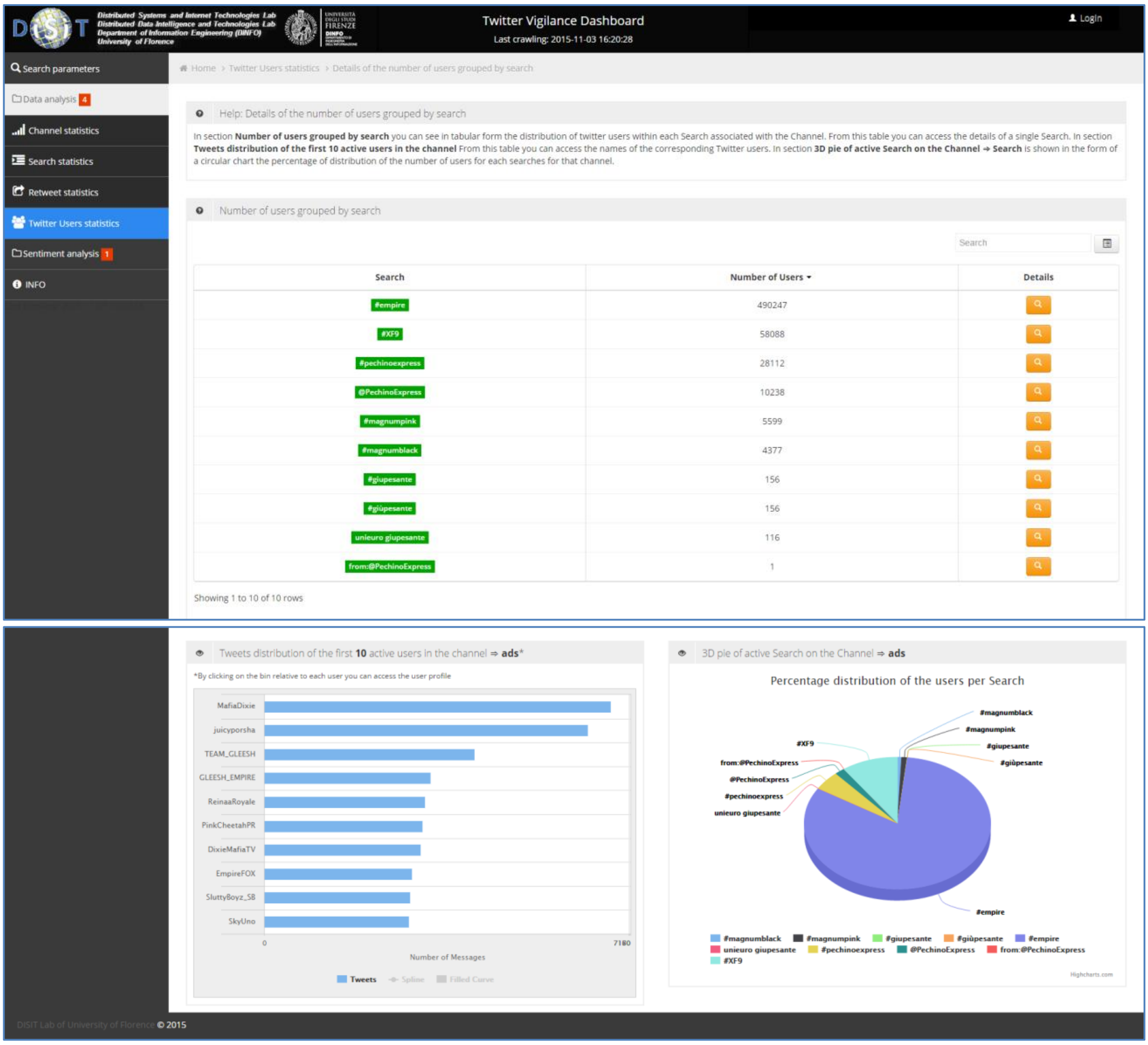
Overview of the number of users per channel

Channel	Number of Users	Details
CALDO	330287	
EXPO2015	189832	
cardboard	189274	
ads	178587	
Firenze	172441	
mymeteo	141753	
NASA New Horizons	137406	
Allertameteo TOSCANA	88358	
nutella	81527	
laudatoSi	75784	

Showing 1 to 10 of 33 rows 10 records per page

Figure 9: Twitter Users statistics

Clicking on the button in the Details column of the table leads to the detail page of the channel, Figure10, where there are two graphs and a table: The table represents the number of messages related to individual searches for that channel, the histogram shows the top 10 users with the greatest number of posts on the channel and the pie chart shows the distribution of users in individual research belonging to the channel.


Distributed Systems and Internet Technologies Lab
Distributed Data Intelligence and Technologies Lab
Department of Information Engineering (DINFO)
University of Florence
Login
Search parameters
Data analysis 4
Channel statistics
Search statistics
Retweet statistics
Twitter Users statistics
Sentiment analysis 1
INFO
Home › Twitter Users statistics › Details of the number of users grouped by search

Twitter Vigilance Dashboard

Last crawling: 2015-11-03 16:20:28

Help: Details of the number of users grouped by search

In section **Number of users grouped by search** you can see in tabular form the distribution of twitter users within each Search associated with the Channel. From this table you can access the details of a single Search. In section **Tweets distribution of the first 10 active users in the channel** From this table you can access the names of the corresponding Twitter users. In section **3D pie of active Search on the Channel ⇒ Search** is shown in the form of a circular chart the percentage of distribution of the number of users for each searches for that channel.

Number of users grouped by search

Search	Number of Users ▾	Details
#empire	490247	🔍
#XF9	58088	🔍
#pechinoexpress	28112	🔍
@PechinoExpress	10238	🔍
#magnumpink	5599	🔍
#magnumblack	4377	🔍
#giupesante	156	🔍
#giùpesante	156	🔍
unieuro giupesante	116	🔍
from:@PechinoExpress	1	🔍

Showing 1 to 10 of 10 rows

Tweets distribution of the first **10** active users in the channel ⇒ **ads***

*By clicking on the bin relative to each user you can access the user profile

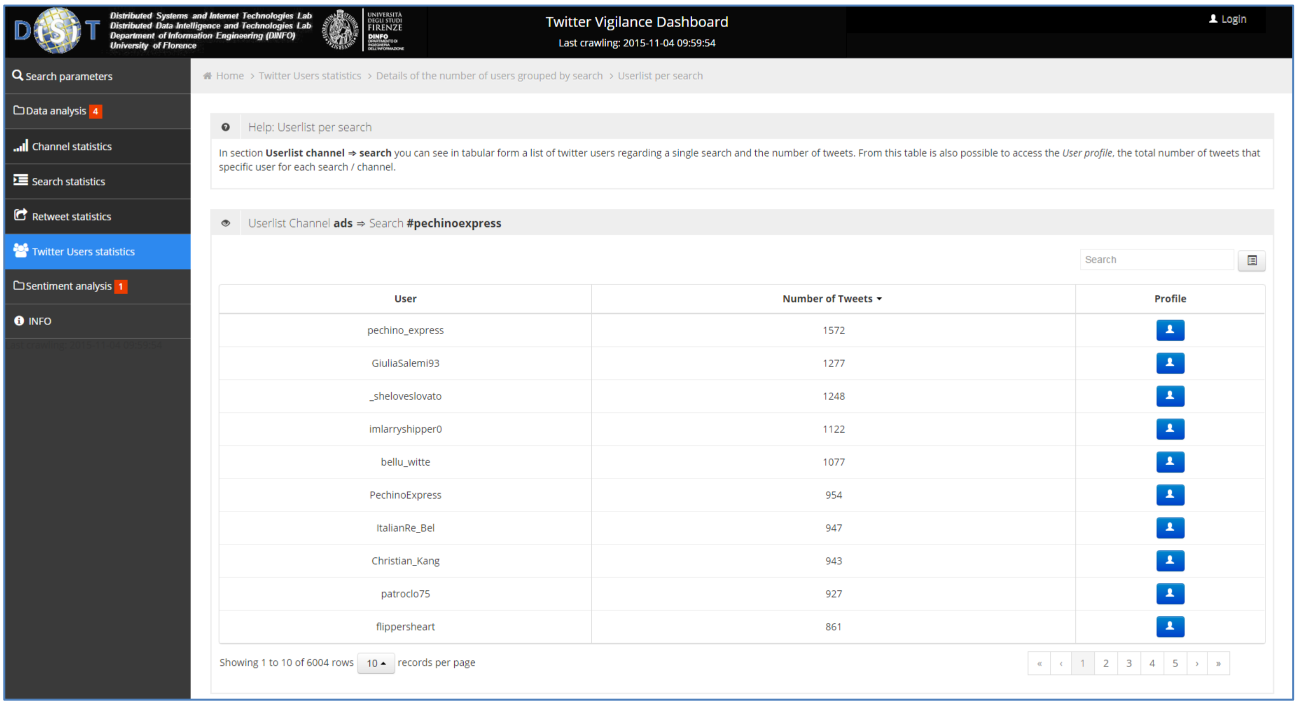
3D pie of active Search on the Channel ⇒ **ads**

DISIT Lab of University of Florence © 2015

Figure 10: Details of the number of users grouped by search

Selecting the bin of a user it's possible to access the profile of the user.

Clicking on the button in the Details column of the table leads to the detail page of the Search, Figure 11, where is displayed a table: the table represents the list of users who have written at least one message among those recovered from the Search and the number of posts written by each user.


Distributed Systems and Internet Technologies Lab
Distributed Data Intelligence and Technologies Lab
Department of Information Engineering (DINFO)
University of Florence
Twitter Vigilance Dashboard
Last crawling: 2015-11-04 09:59:54
Login
Search parameters
Data analysis 4
Channel statistics
Search statistics
Retweet statistics
Twitter Users statistics
Sentiment analysis 1
INFO
Home > Twitter Users statistics > Details of the number of users grouped by search > Userlist per search
Search
« ‹ 1 2 3 4 5 › »

Help: Userlist per search

In section **Userlist channel ⇒ search** you can see in tabular form a list of twitter users regarding a single search and the number of tweets. From this table is also possible to access the *User profile*, the total number of tweets that specific user for each search / channel.

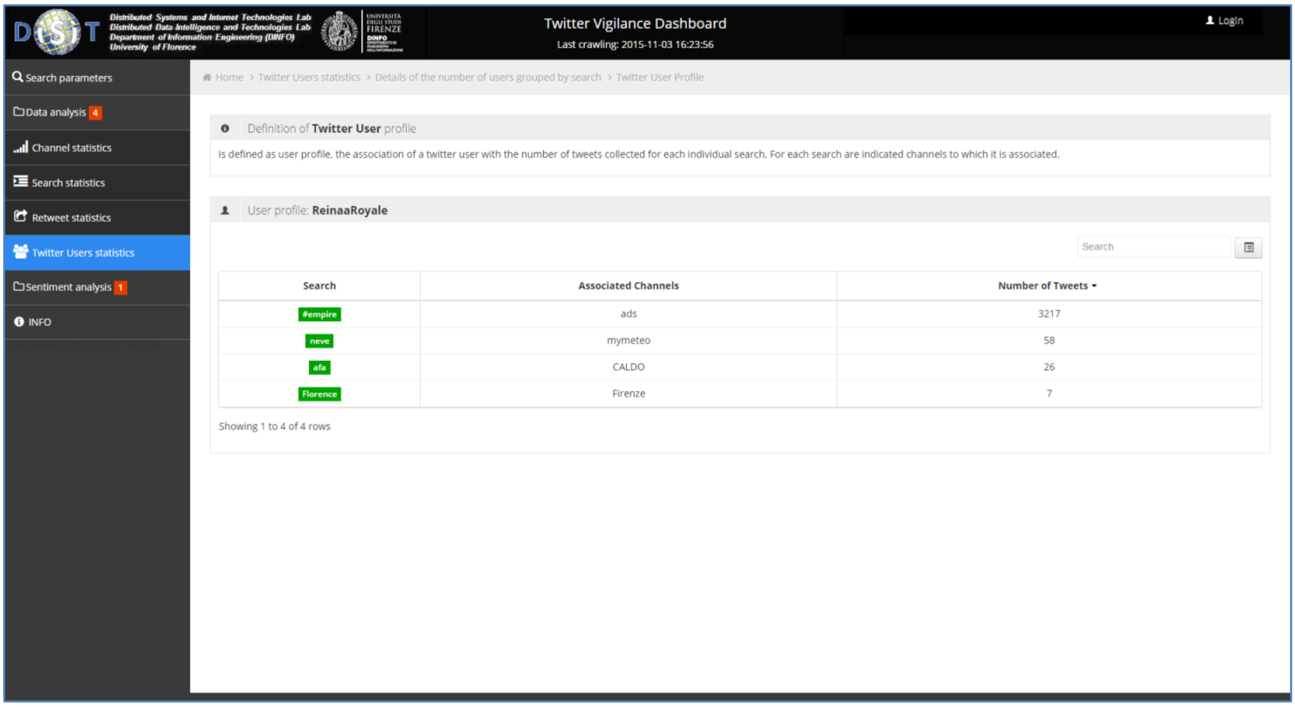
Userlist Channel **ads** ⇒ Search **#pechinoexpress**

User	Number of Tweets ▾	Profile
pechino_express	1572	👤
GiuliaSalemi93	1277	👤
_sheloveslovato	1248	👤
imlarryshipper0	1122	👤
bellu_witte	1077	👤
PechinoExpress	954	👤
ItalianRe_Bel	947	👤
Christian_Kang	943	👤
patroclo75	927	👤
flippersheart	861	👤

Showing 1 to 10 of 6004 rows 10 ▴ records per page

Figure 11: Details of user grouped by Search

Clicking on Profile button it's possible to access the profile of the user (Figure 12).


Distributed Systems and Internet Technologies Lab
Distributed Data Intelligence and Technologies Lab
Department of Information Engineering (DINFO)
University of Florence
Twitter Vigilance Dashboard
Last crawling: 2015-11-03 16:23:56
Login
Search parameters
Data analysis 4
Channel statistics
Search statistics
Retweet statistics
Twitter Users statistics
Sentiment analysis 1
INFO
Home > Twitter Users statistics > Details of the number of users grouped by search > Twitter User Profile
Search

Definition of **Twitter User** profile

Is defined as user profile, the association of a twitter user with the number of tweets collected for each individual search. For each search are indicated channels to which it is associated.

User profile: **ReinaaRoyale**

Search	Associated Channels	Number of Tweets ▾
#empire	ads	3217
neve	mymeteo	58
afa	CALDO	26
Florence	Firenze	7

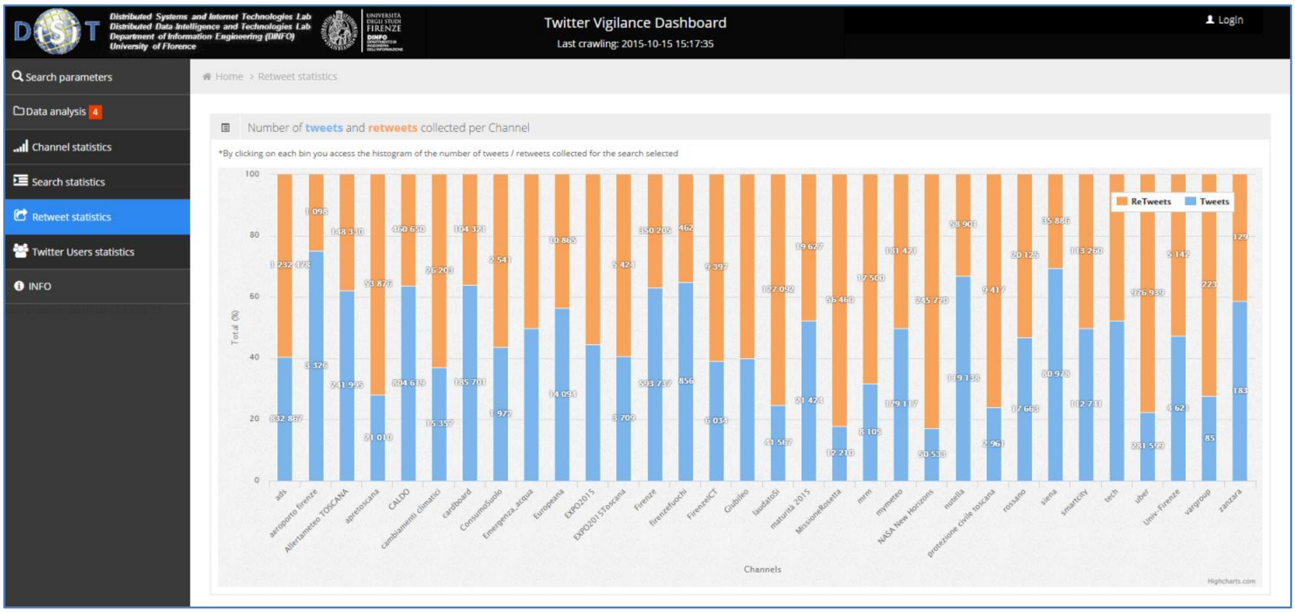
Showing 1 to 4 of 4 rows

Figure 12: User profile

Page represented in Figure 12 displays the association of a twitter user with the number of tweets collected for each individual search. For each search outlines the channels to which it is associated.

3.1.1.4 Retweets

This page displays a bar chart that shows the number of tweets and retweets for each channel.



Twitter Vigilance Dashboard

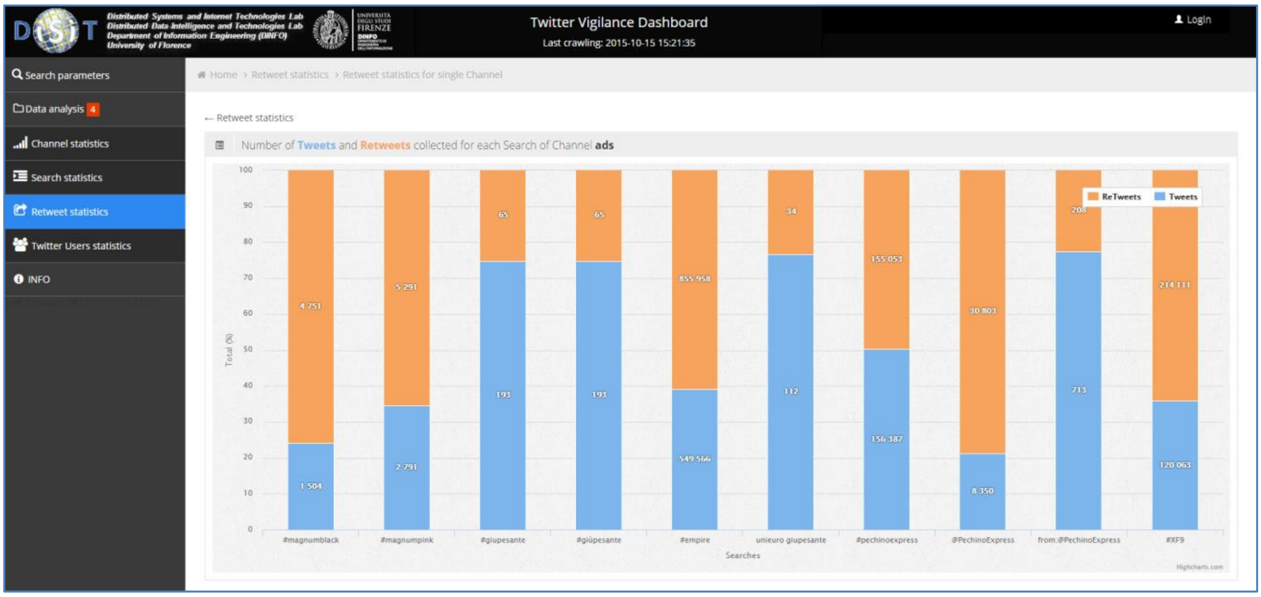
Last crawling: 2015-10-15 15:17:35

Number of tweets and retweets collected per Channel

*By clicking on each bin you access the histogram of the number of tweets / retweets collected for the search selected

Figure 13: Retweet statistics page

Clicking on a single bar is shown the details of the research associated with the selected channel



Twitter Vigilance Dashboard

Last crawling: 2015-10-15 15:21:35

Number of Tweets and Retweets collected for each Search of Channel **ads**

Figure 14: Retweet statistics for single Channel page

3.1.1.5 Administrator Frontend

Respect to the User part, some parts remain unchanged while there are others for the management of the backend of the platform. The pages that are part of the Data Analysis section are identical to those of the User but instead of representing only channels associated to the user, in the case of Administrator are present all channels inserted in the platform.

Even the page "Search parameters" has the same features. As seen in Figure 15 there are two tables: the first shows the channel list, the second the list of searches. In the case of the channel table, the user will only display his channels while the Administrator all those present in the platform. The Search table can be viewed by the Administrator with the ability to edit and delete. The channel table contains buttons to edit, delete and put in standby channels.

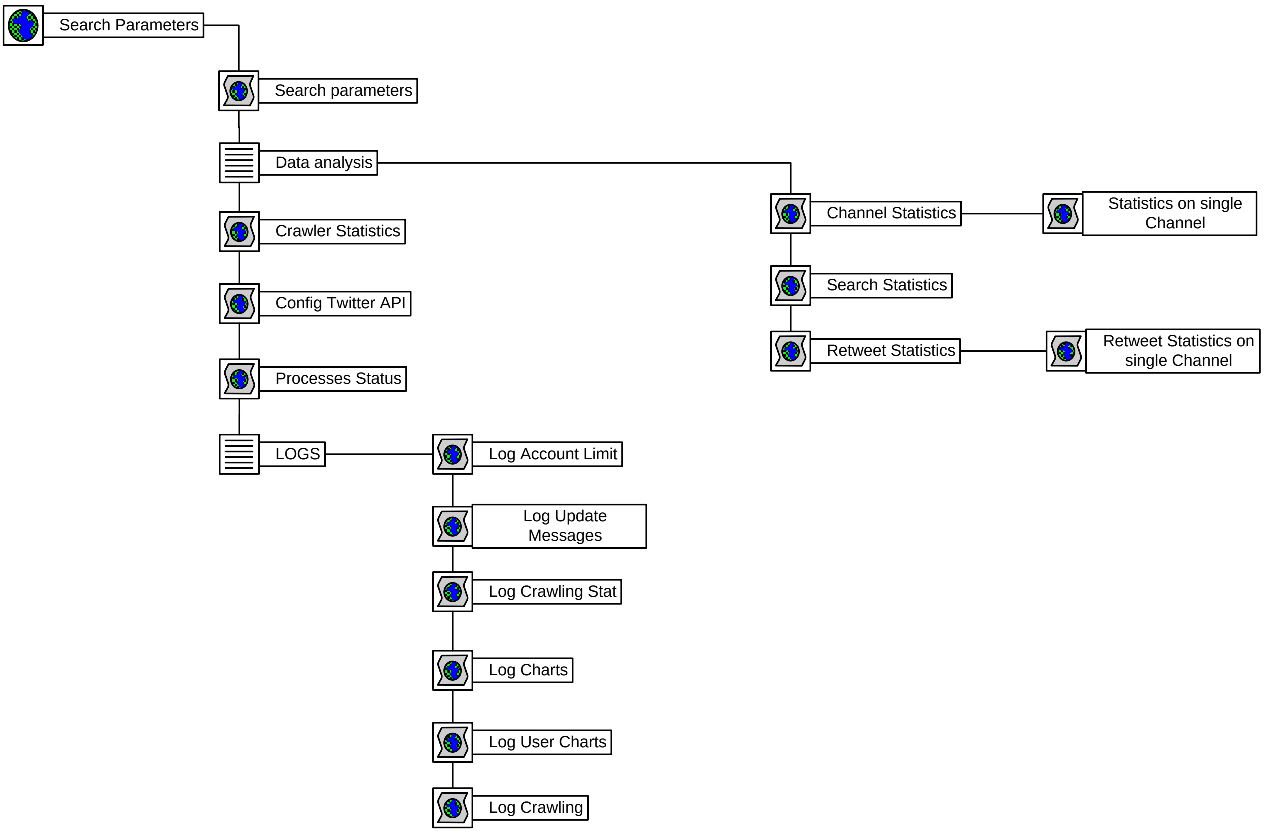
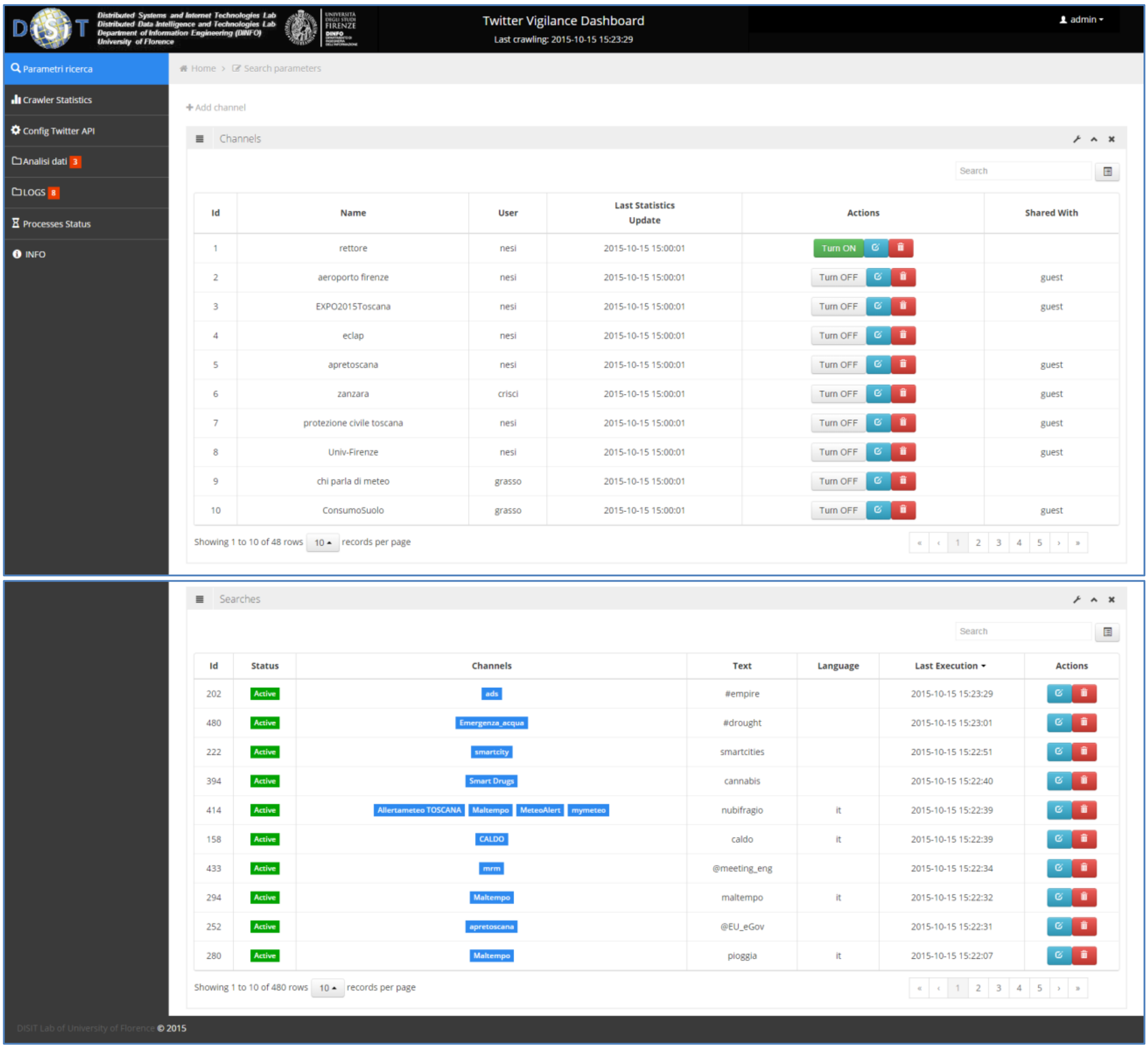


Figure 15: Structure of Administrator Frontend


Distributed Systems and Internet Technologies Lab
Distributed Data Intelligence and Technologies Lab
Department of Information Engineering (DINFO)
University of Florence
UNIVERSITÀ DEGLI STUDI FIRENZE
DINFO
admin
Parametri ricerca
Crawler Statistics
Config Twitter API
Analisi dati 3
LOGS 8
Processes Status
INFO
Home > Search parameters
+ Add channel
Search

Twitter Vigilance Dashboard

Last crawling: 2015-10-15 15:23:29

Channels

Id	Name	User	Last Statistics Update	Actions	Shared With
1	rettore	nesi	2015-10-15 15:00:01	Turn ON	
2	aeroporto firenze	nesi	2015-10-15 15:00:01	Turn OFF	guest
3	EXPO2015Toscana	nesi	2015-10-15 15:00:01	Turn OFF	guest
4	eclap	nesi	2015-10-15 15:00:01	Turn OFF	
5	apretoscana	nesi	2015-10-15 15:00:01	Turn OFF	guest
6	zanzara	crisci	2015-10-15 15:00:01	Turn OFF	guest
7	protezione civile toscana	nesi	2015-10-15 15:00:01	Turn OFF	guest
8	Univ-Firenze	nesi	2015-10-15 15:00:01	Turn OFF	guest
9	chi parla di meteo	grasso	2015-10-15 15:00:01	Turn OFF	
10	ConsumoSuolo	grasso	2015-10-15 15:00:01	Turn OFF	guest

Showing 1 to 10 of 48 rows 10 records per page

« ‹ 1 2 3 4 5 › »

Searches

Id	Status	Channels	Text	Language	Last Execution	Actions
202	Active	ads	#empire		2015-10-15 15:23:29	
480	Active	Emergenza_acqua	#drought		2015-10-15 15:23:01	
222	Active	smartcity	smartcities		2015-10-15 15:22:51	
394	Active	Smart Drugs	cannabis		2015-10-15 15:22:40	
414	Active	Allertameteo TOSCANA Maltempo MeteoAlert mymeteo	nubifragio	it	2015-10-15 15:22:39	
158	Active	CALDO	caldo	it	2015-10-15 15:22:39	
433	Active	mrm	@meeting_eng		2015-10-15 15:22:34	
294	Active	Maltempo	maltempo	it	2015-10-15 15:22:32	
252	Active	apretoscana	@EU_eGov		2015-10-15 15:22:31	
280	Active	Maltempo	pioggia	it	2015-10-15 15:22:07	

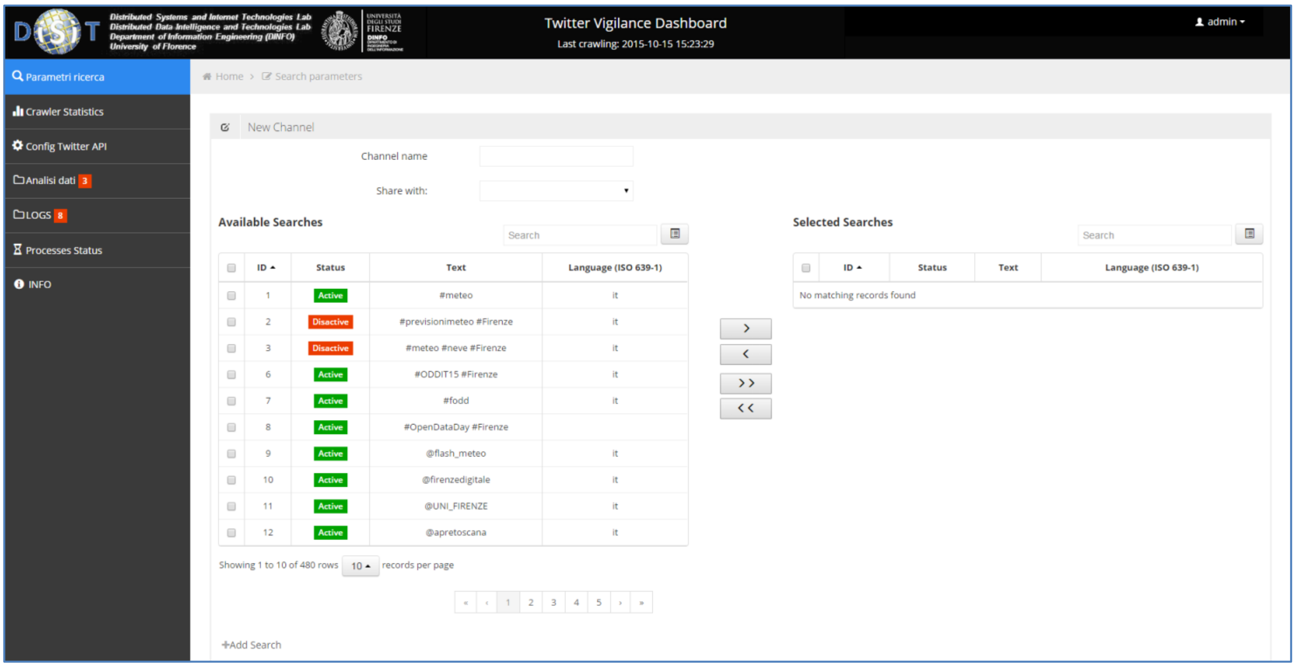
Showing 1 to 10 of 480 rows 10 records per page

« ‹ 1 2 3 4 5 › »

DISIT Lab of University of Florence © 2015

Figure 16: Search parameters page

Via the link "Add Channel" opens a section that allows you to enter a new channel (Figure 17): required fields are the name of the channel and the list of the searches to associate with. There is the possibility to publish the channel sharing it with the "guest" user. Also via the link "Add Search" leads to a form for entering a new search that is automatically assigned to the new channel. The input section of a new channel has the same characteristics as that for editing of the channel.


Distributed Systems and Internet Technologies Lab
Distributed Data Intelligence and Technologies Lab
Department of Information Engineering (DINFO)
University of Florence
Università degli Studi Firenze
Twitter Vigilance Dashboard
Last crawling: 2015-10-15 15:23:29
admin
Parametri ricerca
Crawler Statistics
Config Twitter API
Analisi dati 3
LOGS 8
Processes Status
INFO
Home > Search parameters
New Channel
Channel name
Share with:
Search
>
<
>>
<<
+Add Search

Available Searches

□	ID ▲	Status	Text	Language (ISO 639-1)
□	1	Active	#meteo	it
□	2	Disactive	#previsionimeteo #Firenze	it
□	3	Disactive	#meteo #neve #Firenze	it
□	6	Active	#ODDIT15 #Firenze	it
□	7	Active	#fodd	it
□	8	Active	#OpenDataDay #Firenze	
□	9	Active	@flash_meteo	it
□	10	Active	@firenzedigitale	it
□	11	Active	@UNI_FIRENZE	it
□	12	Active	@apretoscana	it

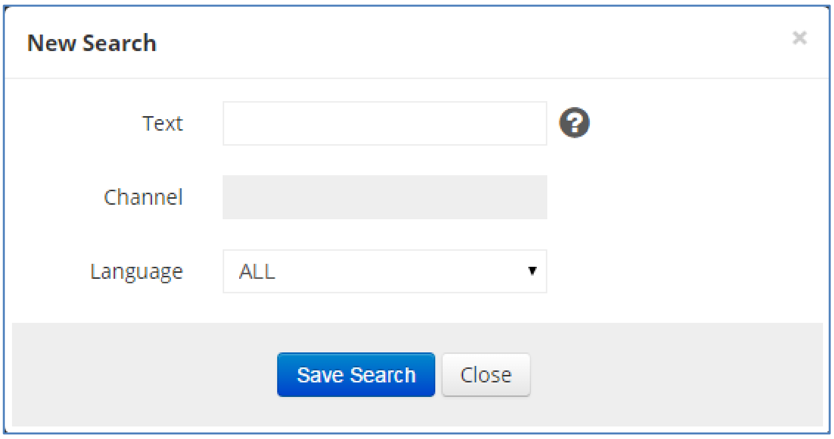
Showing 1 to 10 of 480 rows 10 records per page

« ‹ 1 2 3 4 5 › »

Selected Searches

□	ID ▲	Status	Text	Language (ISO 639-1)
No matching records found				

Figure 17: Search parameters page -> add new channel



Text:

Channel:

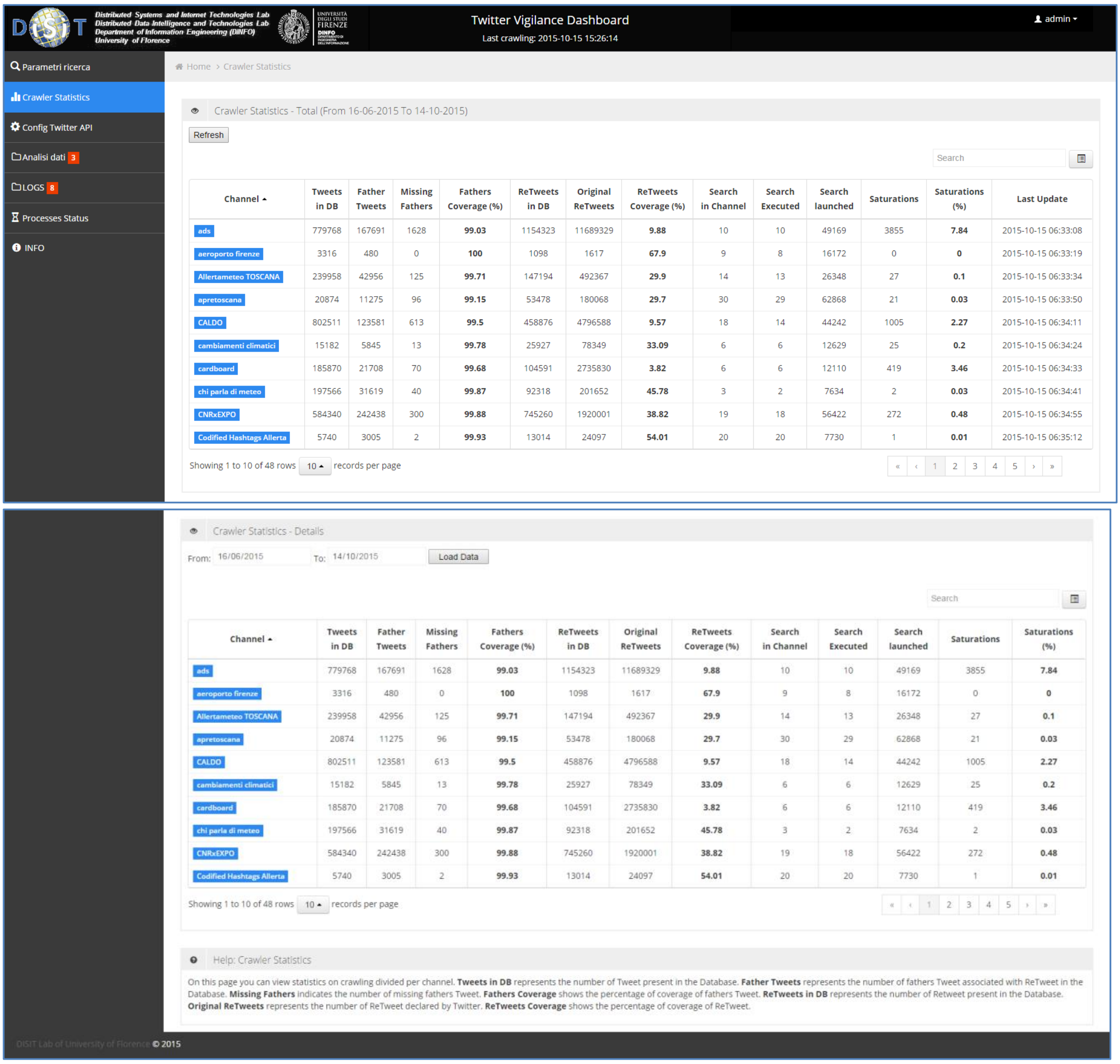
Language: ALL

Figure 18: Add new Search form

In Figure 19 is shown the page with the statistics of crawling: there are two tables, one for the total summary of the statistics and the daily details. The statistics shown are divided per channel and include:

- Number of Tweets
- Number of fathers Tweets
- Number of missing fathers Tweets
- Coverage of fathers Tweet

- Number of Retweets
- Number of Retweets declared by Twitter
- Coverage of Retweets
- Number of searches for that channel
- Number of searches performed
- Number of requests made to Twitter
- Number of saturations
- Percentage of Saturations


DISIT
Distributed Systems and Internet Technologies Lab
Distributed Data Intelligence and Technologies Lab
Department of Information Engineering (DINFO)
University of Florence
UNIVERSITÀ DEGLI STUDI FIRENZE
DINFO
admin
Parametri ricerca
Crawler Statistics
Config Twitter API
Analisi dati 3
LOGS 8
Processes Status
INFO
Home > Crawler Statistics
Refresh
Search
Load Data
« ‹ 1 2 3 4 5 › »

Twitter Vigilance Dashboard

Last crawling: 2015-10-15 15:26:14

Crawler Statistics - Total (From 16-06-2015 To 14-10-2015)

Channel ▲	Tweets in DB	Father Tweets	Missing Fathers	Fathers Coverage (%)	ReTweets in DB	Original ReTweets	ReTweets Coverage (%)	Search in Channel	Search Executed	Search launched	Saturations	Saturations (%)	Last Update
ads	779768	167691	1628	**99.03**	1154323	11689329	**9.88**	10	10	49169	3855	**7.84**	2015-10-15 06:33:08
aeroporto firenze	3316	480	0	**100**	1098	1617	**67.9**	9	8	16172	0	**0**	2015-10-15 06:33:19
Allertameteo TOSCANA	239958	42956	125	**99.71**	147194	492367	**29.9**	14	13	26348	27	**0.1**	2015-10-15 06:33:34
apretoscana	20874	11275	96	**99.15**	53478	180068	**29.7**	30	29	62868	21	**0.03**	2015-10-15 06:33:50
CALDO	802511	123581	613	**99.5**	458876	4796588	**9.57**	18	14	44242	1005	**2.27**	2015-10-15 06:34:11
cambiamenti climatici	15182	5845	13	**99.78**	25927	78349	**33.09**	6	6	12629	25	**0.2**	2015-10-15 06:34:24
cardboard	185870	21708	70	**99.68**	104591	2735830	**3.82**	6	6	12110	419	**3.46**	2015-10-15 06:34:33
chi parla di meteo	197566	31619	40	**99.87**	92318	201652	**45.78**	3	2	7634	2	**0.03**	2015-10-15 06:34:41
CNRxEXPO	584340	242438	300	**99.88**	745260	1920001	**38.82**	19	18	56422	272	**0.48**	2015-10-15 06:34:55
Codified Hashtags Allerta	5740	3005	2	**99.93**	13014	24097	**54.01**	20	20	7730	1	**0.01**	2015-10-15 06:35:12

Showing 1 to 10 of 48 rows 10 ▲ records per page

Crawler Statistics - Details

From: 16/06/2015 To: 14/10/2015

Channel ▲	Tweets in DB	Father Tweets	Missing Fathers	Fathers Coverage (%)	ReTweets in DB	Original ReTweets	ReTweets Coverage (%)	Search in Channel	Search Executed	Search launched	Saturations	Saturations (%)
ads	779768	167691	1628	**99.03**	1154323	11689329	**9.88**	10	10	49169	3855	**7.84**
aeroporto firenze	3316	480	0	**100**	1098	1617	**67.9**	9	8	16172	0	**0**
Allertameteo TOSCANA	239958	42956	125	**99.71**	147194	492367	**29.9**	14	13	26348	27	**0.1**
apretoscana	20874	11275	96	**99.15**	53478	180068	**29.7**	30	29	62868	21	**0.03**
CALDO	802511	123581	613	**99.5**	458876	4796588	**9.57**	18	14	44242	1005	**2.27**
cambiamenti climatici	15182	5845	13	**99.78**	25927	78349	**33.09**	6	6	12629	25	**0.2**
cardboard	185870	21708	70	**99.68**	104591	2735830	**3.82**	6	6	12110	419	**3.46**
chi parla di meteo	197566	31619	40	**99.87**	92318	201652	**45.78**	3	2	7634	2	**0.03**
CNRxEXPO	584340	242438	300	**99.88**	745260	1920001	**38.82**	19	18	56422	272	**0.48**
Codified Hashtags Allerta	5740	3005	2	**99.93**	13014	24097	**54.01**	20	20	7730	1	**0.01**

Showing 1 to 10 of 48 rows 10 ▲ records per page

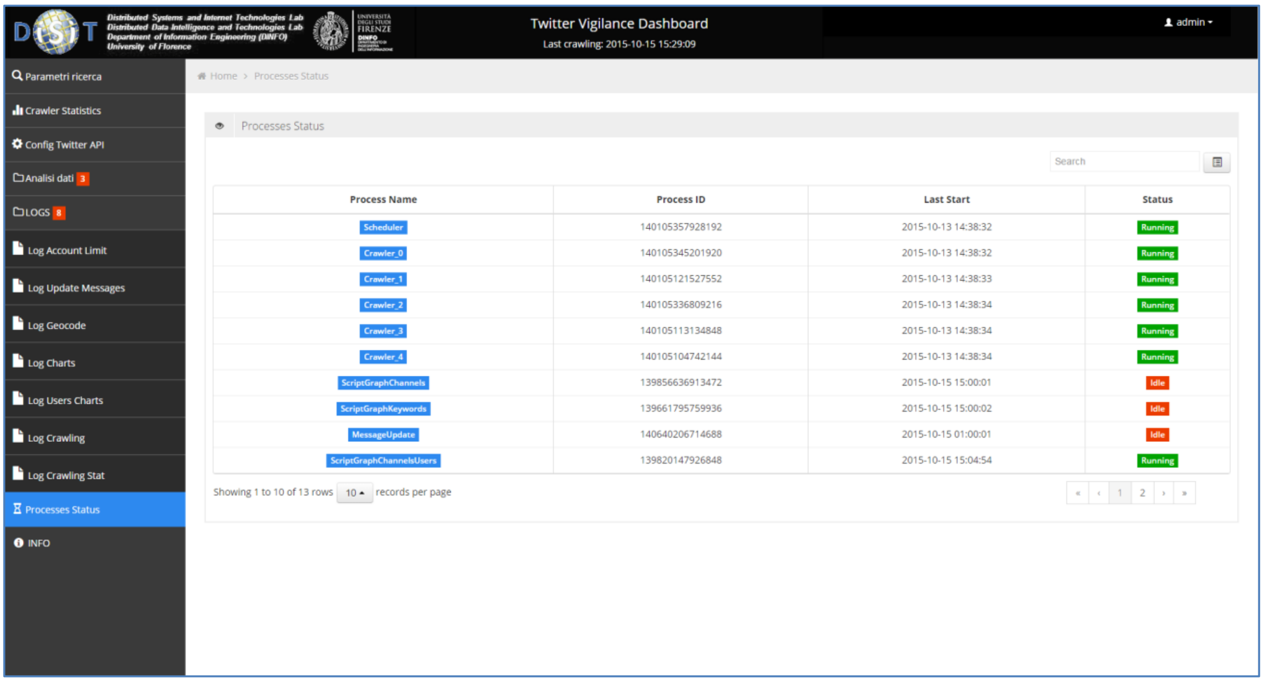
Help: Crawler Statistics

On this page you can view statistics on crawling divided per channel. **Tweets in DB** represents the number of Tweet present in the Database. **Father Tweets** represents the number of fathers Tweet associated with ReTweet in the Database. **Missing Fathers** indicates the number of missing fathers Tweet. **Fathers Coverage** shows the percentage of coverage of fathers Tweet. **ReTweets in DB** represents the number of Retweet present in the Database. **Original ReTweets** represents the number of ReTweet declared by Twitter. **ReTweets Coverage** shows the percentage of coverage of ReTweet.

DISIT Lab of University of Florence © 2015

Figure 19: Crawler statistics page

In Figure 20 all active processes on the backend and their status are listed: in particular displays the job status (Running, Idle), the process ID and the date and time of the execution of the process.



Twitter Vigilance Dashboard

Last crawling: 2015-10-15 15:29:09

Processes Status

Process Name	Process ID	Last Start	Status
Scheduler	140105357928192	2015-10-13 14:38:32	Running
Crawler_0	140105345201920	2015-10-13 14:38:32	Running
Crawler_1	140105121527552	2015-10-13 14:38:33	Running
Crawler_2	140105336809216	2015-10-13 14:38:34	Running
Crawler_3	140105113134848	2015-10-13 14:38:34	Running
Crawler_4	140105104742144	2015-10-13 14:38:34	Running
ScriptGraphChannels	139856636913472	2015-10-15 15:00:01	Idle
ScriptGraphKeywords	139661795759936	2015-10-15 15:00:02	Idle
MessageUpdate	140640206714688	2015-10-15 01:00:01	Idle
ScriptGraphChannelsUsers	139820147926848	2015-10-15 15:04:54	Running

Showing 1 to 10 of 13 rows 10 records per page

Figure 20: Processes status page

4 Bibliografia

[1] M. Faheem e P. Senellart, «Intelligent and Adaptive Crawling of Web Applications for Web Archiving,» in *13th International Conference, ICWE Proceedings*, Aalborg, Denmark, 2013.

[2] C.-I. Wong, K.-Y. Wong, K.-W. NG, W. Fan e K.-H. Yeung, «Design of a Crawler for Online Social Networks Analysis,» *WSEAS Transactions on Communications,* vol. 13, p. 263, 2014.

[3] House of Kaizen, «SOCIAL MEDIA MONITORING TOOL BUYER'S GUIDE».

[4] B. Batrinca e P. C. Treleaven, «Social media analytics: a survey of techniques, tools and platforms,» *AI & SOCIETY,* vol. 30, pp. 89-116, 2015.

[5] «80legs,» [Online]. Available: http://80legs.com/.

[6] Sequentum, «http://www.visualwebripper.com/Default.aspx,» [Online].

[7] «Helium Scraper,» [Online]. Available: http://www.heliumscraper.com/en/index.php?p=home.

[8] «Prompt Cloud,» [Online]. Available: http://www.promptcloud.com/.

[9] M. Castellanos, M. Hsu, U. Dayal, R. Ghosh, M. Dekhil, C. Ceja, M. Puchi e P. Ruiz, «Intention insider: discovering people's intentions in the social channel,» in *Proceedings of the 15th International Conference on Extending Database Technology*, Berlin, Germany, 2012.

[10] «Open Amplify,» [Online]. Available: http://www.openamplify.com/.

[11] «Clarabridge,» [Online]. Available: http://clarabridge.com/.

[12] «BrandWatch,» [Online]. Available: http://www.brandwatch.com/.

[13] «Opinion Crawl,» [Online]. Available: http://www.opinioncrawl.com/.

[14] «Social Report,» [Online]. Available: http://www.socialreport.com/.

[15] «Mozenda,» [Online]. Available: http://www.mozenda.com/.

[16] «beevolve,» [Online]. Available: http://www.beevolve.com/.

[17] «Meltwater,» [Online]. Available: http://www.meltwater.com/.

[18] «Viralheat,» [Online]. Available: https://www.viralheat.com/.

[19] «SAS Sentiment Analysis,» [Online]. Available: http://www.sas.com/en_us/software/analytics/sentiment-analysis.html.

[20] «Dataminr,» [Online]. Available: https://www.dataminr.com/.

[21] «tracx,» [Online]. Available: http://www.tracx.com/.

[22] «Royalty,» [Online]. Available: http://www.roialty.com/.

[23] A. Talamo, *Progettazione e sviluppo di un modulo per l'integrazione di Twitter nel CMS Drupal e analisi descrittiva del flusso dei Tweets,* Firenze, 2015.

[24] «Twopcharts, sito che fornisce statistiche ufficiali su twitter,» [Online]. Available: http://twopcharts.com.

[25] «Supporto Twitter,» [Online]. Available: https://support.twitter.com/articles/119138-types-of-tweets-and-where-they-appear.