



UNIVERSITÀ
DEGLI STUDI
FIRENZE

DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

DISIT
DISTRIBUTED SYSTEMS
AND INTERNET
TECHNOLOGIES LAB



Twitter Vigilance: Modelli e Strumenti per l'Analisi e lo Studio di Dati Social Media ed il Monitoraggio in Real Time

Daniele Cenni, Paolo Nesi, Gianni Pantaleo, Irene Paoli, Imad Zaza

DISIT Lab, Distributed [Systems and internet | Data Intelligence
and] Technologies Lab, Dep. of Information Engineering (DINFO),

University of Florence, Italy, Fax: 0039-055-2758570

<http://www.disit.dinfo.unifi.it>, paolo.nesi@unifi.it

<http://www.disit.org/tv>

<http://www.disit.org/rttv>

Twitter Vigilance



*Big Data Analytics and Technologies
for Management
Master*
UNIVERSITÀ
DEGLI STUDI
FIRENZE



- **Social media analysis** is becoming a very important instrument to monitor communities, users' preferences, and to make predictions
- **Twitter** is one of the most widespread microblogs allowing users to have a personal news feed and followers attached to it
 - 350,000 Tweets are sent every minute
 - In 2016, Twitter has 310 million monthly active users
(almost the same as the U.S. population)
 - A total of 1.3 billion accounts



Twitter Vigilance: le analisi

- **Analisi e caratterizzazione della comunicazione**
- **Percezione sociale, eventi pubblici, naturali..**
- **Scoprire, identificare e calcolare**
 - **Nascita / crescita di nuove occorrenze** in tempo reale: eventi, fatti, meteo, condizioni critiche, etc.
 - *Supporto alla decisioni, ridurre i tempi di reazione, valutare la percezione, ridurre i costi, incrementare la resilienza come capacità di reagire, diagnosi precoce*
 - **Chi influenza** la comunicazione, le comunità e come: i pusher, gli attori, i *follower*, le sorgenti, etc.
 - **Predizione su eventi** periodici, per esempio presenze ad eventi, presenze sui canali televisivi, vendite aziende, etc.
 - **Misure indirette** basate sulla popolazione: rischio sicurezza, degrado, neve, grandine, vento, fallimenti, etc.



UNIVERSITÀ
DEGLI STUDI
FIRENZE

DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

DISIT
DISTRIBUTED SYSTEMS
AND INTERNET
TECHNOLOGIES LAB

Http://www.disit.org/tv



UNIVERSITÀ
DEGLI STUDI
FIRENZE
DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

Twitter Vigilance Dashboard

Last crawling: 2016-06-23 00:13:01

admin ▾

Search parameters

Crawler Statistics

Config Twitter API

Data analysis 3

Channel Statistics

Search statistics

Retweet Statistics

Twitter Users
statistics

Sentiment analysis 1

LOGS 9

Processes

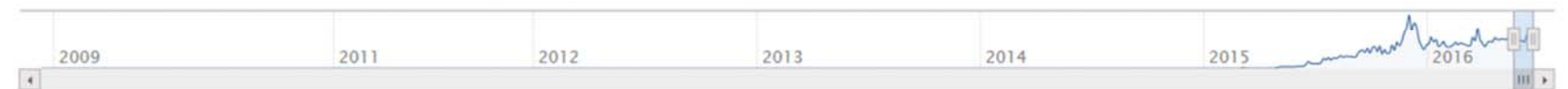
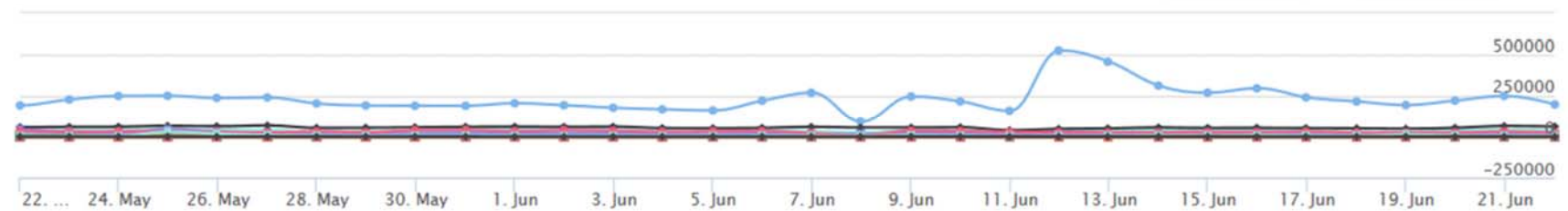
INFO

Home > Channel statistics

Global view of user channels

Zoom 1m 3m 6m YTD 1y All

From May 22, 2016 To Jun 22, 2016



- ads
- aeroporto firenze
- Agenzie-ambientali
- Allertameteo
- Allertameteo TOSCANA
- Alluvione
- apretoscana
- ARPAT
- CALDO
- cambiamenti climatici
- cardboard
- chi parla di meteo
- Clima
- CNRxEXPO
- Codified Hashtags Allerta
- comfortUS
- ConsumoSuolo
- eclap
- Emergenza_acqua
- estra
- Europeana
- EXPO2015
- EXPO2015Toscana
- fallbegins
- Farmacovigilanza
- Firenze
- firenzefuochi
- FirenzeICT
- geocomuniTOS
- Giubileo
- idrogeologico
- iene
- influenza
- iononrischio
- LaMMA
- laudatoSi
- laudatoSi_1
- majorcities2016
- Maltempo
- Maltempo Firenze
- Maltempo Grosseto
- maturità 2015
- MeteoAlert
- MeteoUSER
- MissioneRosetta
- monitAllergie
- mrm
- mugnone2016
- mymeteo
- NASA New Horizons
- nutella
- PAAlert
- PAEnvironment
- PAMeteoNews
- PAMeteoNewsStud
- paolonesi
- papafrancesco
- PAProtCivile
- PA_social_PA
- protezione civile toscana
- reflusso
- resilienza
- rettore
- rossano
- siena
- Smart Drugs
- smartcity
- smartcitybigdata2015
- SOLO allertameteoTOSCANA
- tech
- terrorismo
- TPL
- turismo Firenze
- turismo Firenze_1
- turismo Firenze_2
- uber
- Univ-Firenze
- vargroup
- Vda
- VoragineFi
- Xfactor9
- Xylella
- zanzara
- Events

Hide All

Highcharts.com



Real time Twitter Vigilance

[Http://www.disit.org/rttv](http://www.disit.org/rttv)





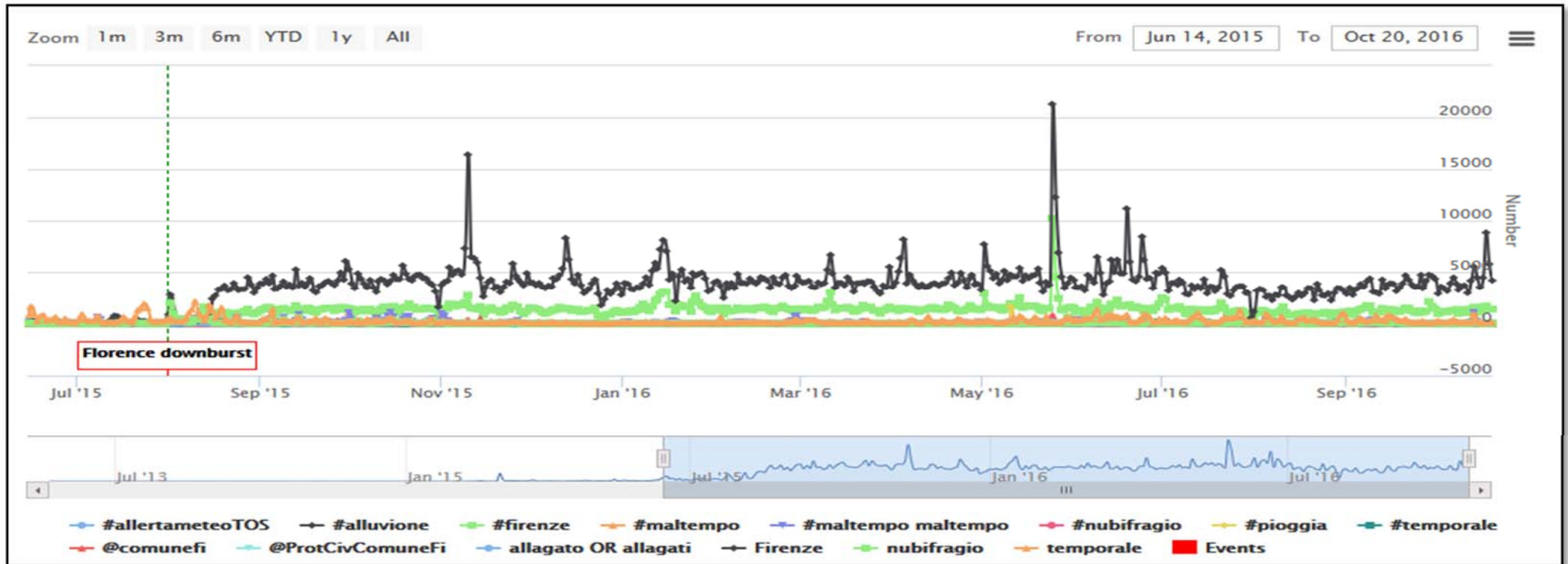
Prediction/Assessment

- Football game results as related to the volume of Tweets
- Number of votes on political elections, via sentiment analysis, SA
- Size and inception of contagious diseases
- marketability of consumer goods
- public health seasonal flu
- box-office revenues for movies
- places to be visited, most visited
- number of people in locations like airports
- audience of TV programmes, political TV shows
- weather forecast information
- Appreciation of services



- Canale “Maltempo Firenze”
- 2 Milioni di Tweet

#allertameteoTOS
 #alluvione | #fiorenze
 #maltempo
 #maltempo maltempo
 #nubifragio | #pioggia
 #temporale | @comunefi
 @ProtCivComuneFi
 allagato OR allagati | Firenze
 nubifragio | temporale



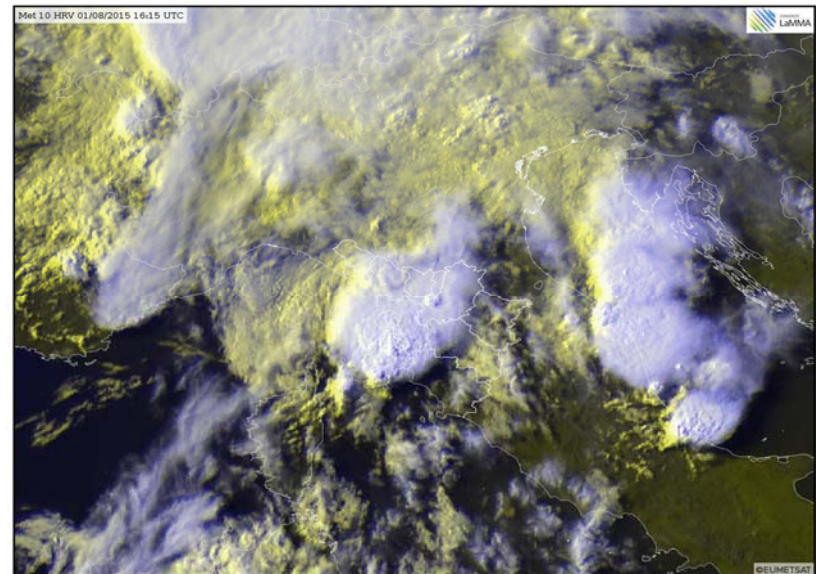
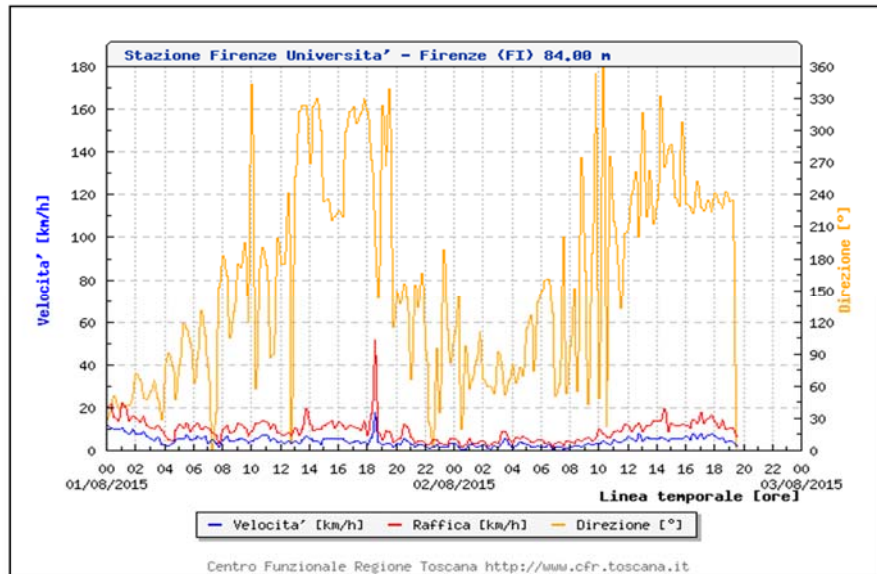
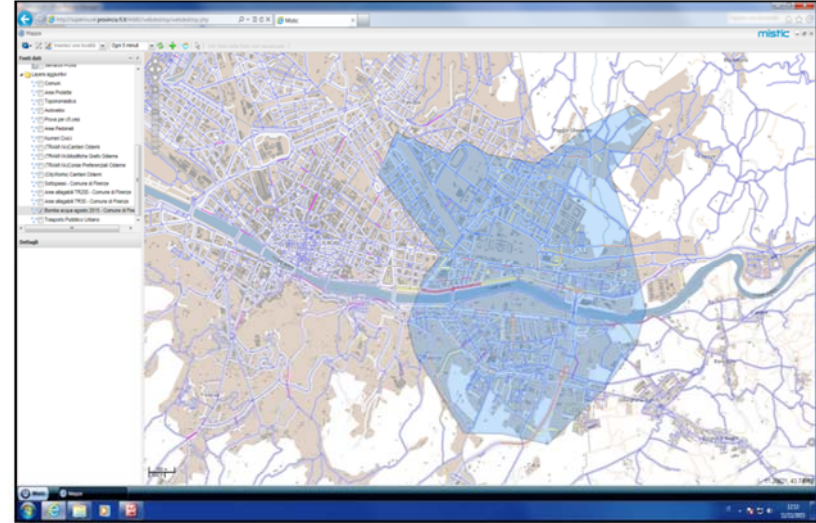
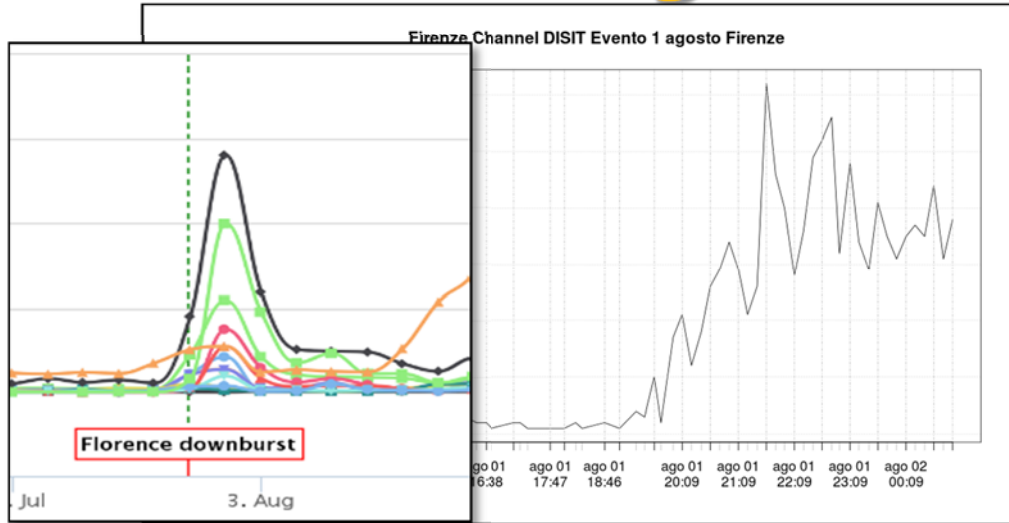


UNIVERSITÀ
DEGLI STUDI
FIRENZE

DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

DISIT
DISTRIBUTED SYSTEMS
AND INTERNET
TECHNOLOGIES LAB

Early Warning Twitter Vigilance and Water Bomb



Twitter Vigilance



Twitter Metrics

- TW: Number of Tweets per **Search/Channel** (as called Volume) , per day, per hour
- RTW: Number of ReTweets per **Search/Channel**, per day, per hour
- NRT/TW: ratio from ReTweets and Tweets per **Search/Channel**, per day, per hour
- NumSearch: number of Tweets including the Search per **Channel**, per day, per hour
- Sentiment Analysis Score per **Search/Channel**, per day, per hour
- Num of xxxxx



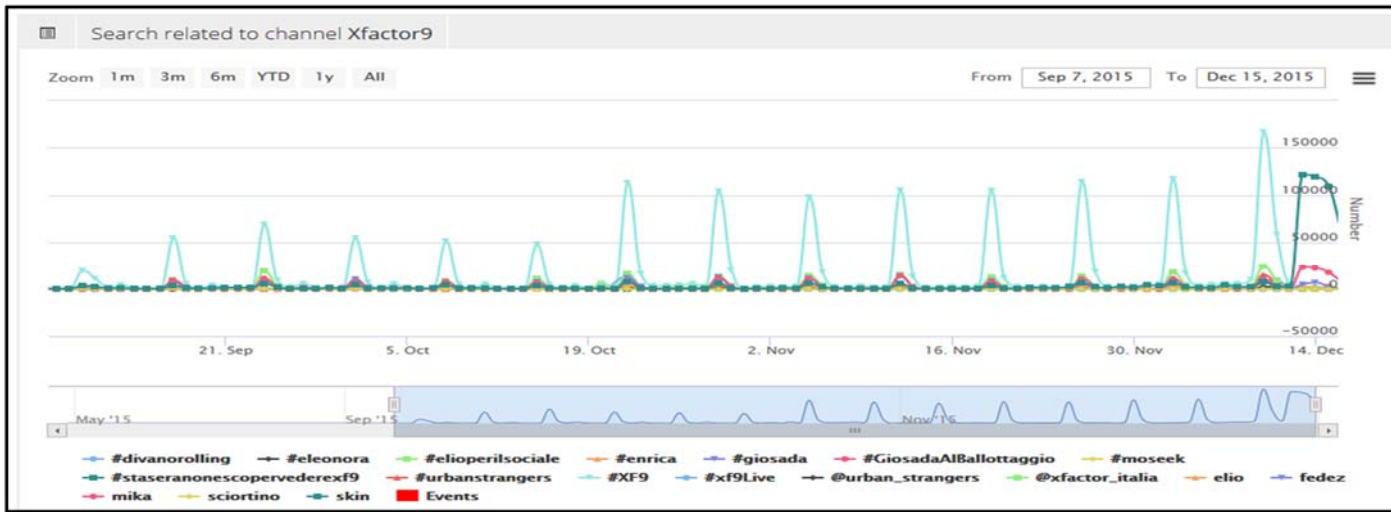
UNIVERSITÀ
DEGLI STUDI
FIRENZE

DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

DISIT
DISTRIBUTED SYSTEMS
AND INTERNET
TECHNOLOGIES LAB

Canale "XFactor9", "XF10"

- #divanorolling
- #eleonora
- #elioperilsociale
- #enrica
- #giosada
- #GiosadaAlBallottaggio
- #moseek
- #staseranonescopervederexf9
- #urbanstrangers
- #XF9
- #xf9Live
- @urban_strangers
- @xfactor_italia
- elio
- fedez
- mika
- sciortino
- skin



- 2 Milioni di Tweet

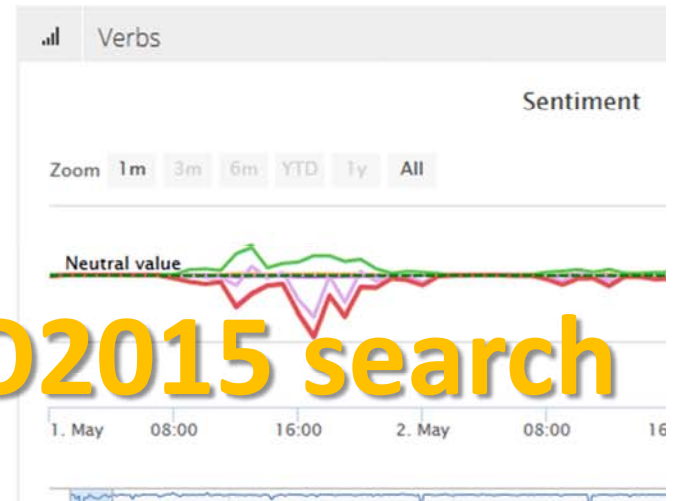
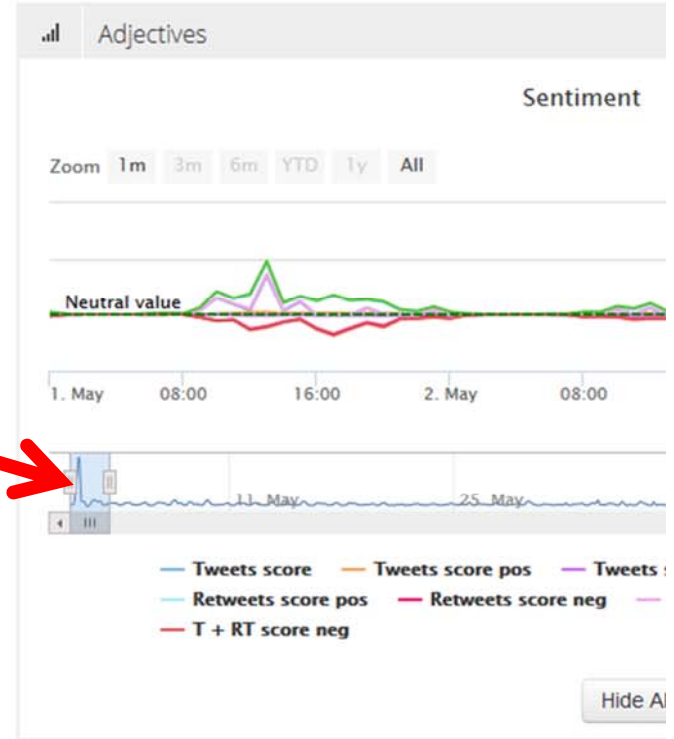
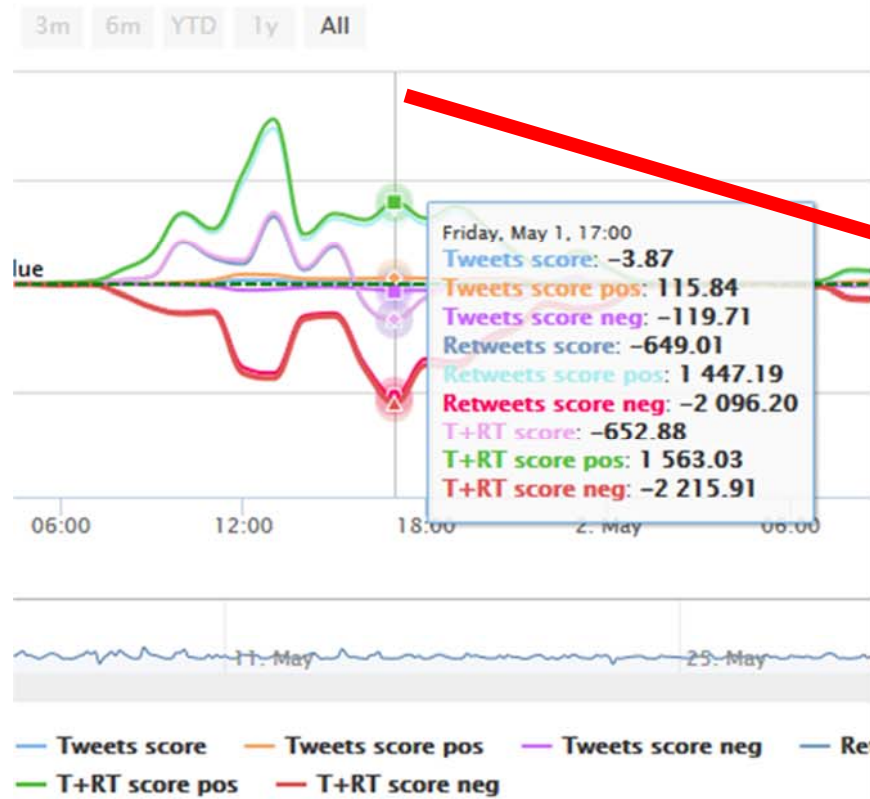




NLP and SA analysis

- Analysis of Twitter text taking into account the specific grammatical constructs and lexicon of twitter
- Extraction of: verbs, adjectives, keywords, citations, hashtags, etc....
- Construction of metrics counting those elements
- Understanding the meaning of the text and taking into account the discourse, the context; disambiguating statements

Sentiment analysis: #Expo2015

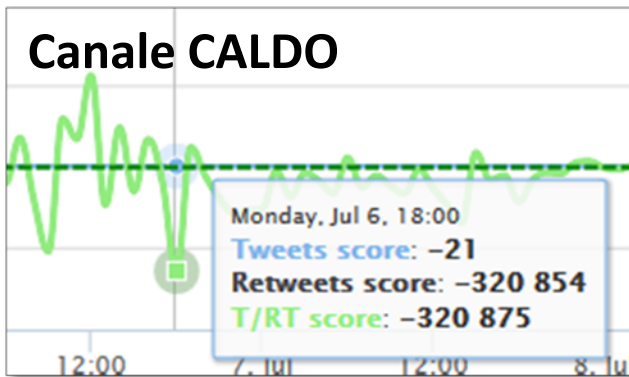


Some views on details of the SA of a Search

Sentiment Analysis on #EXPO2015 search

Sentiment Analysis

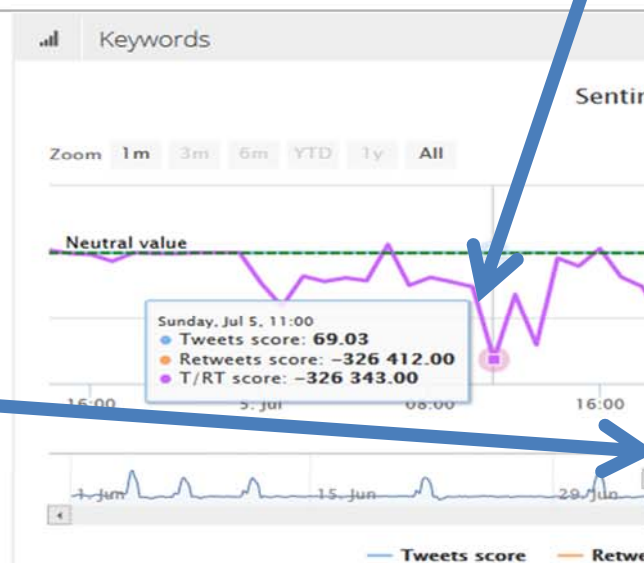
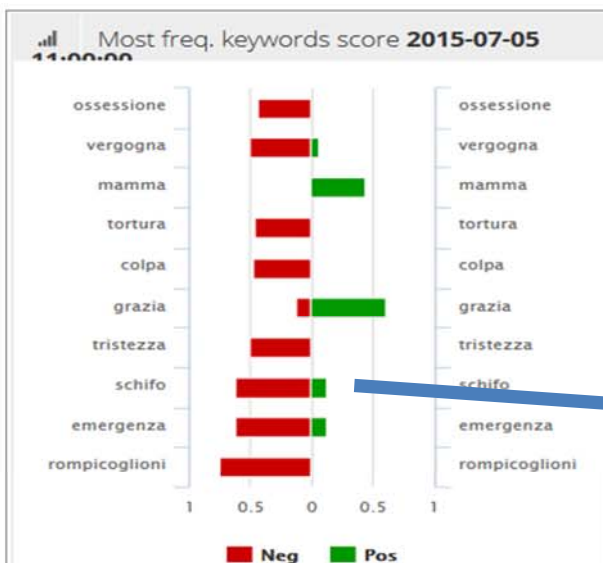
SA sul canale CALDO: la ricerca "caldo" ha dato un sentiment negativo



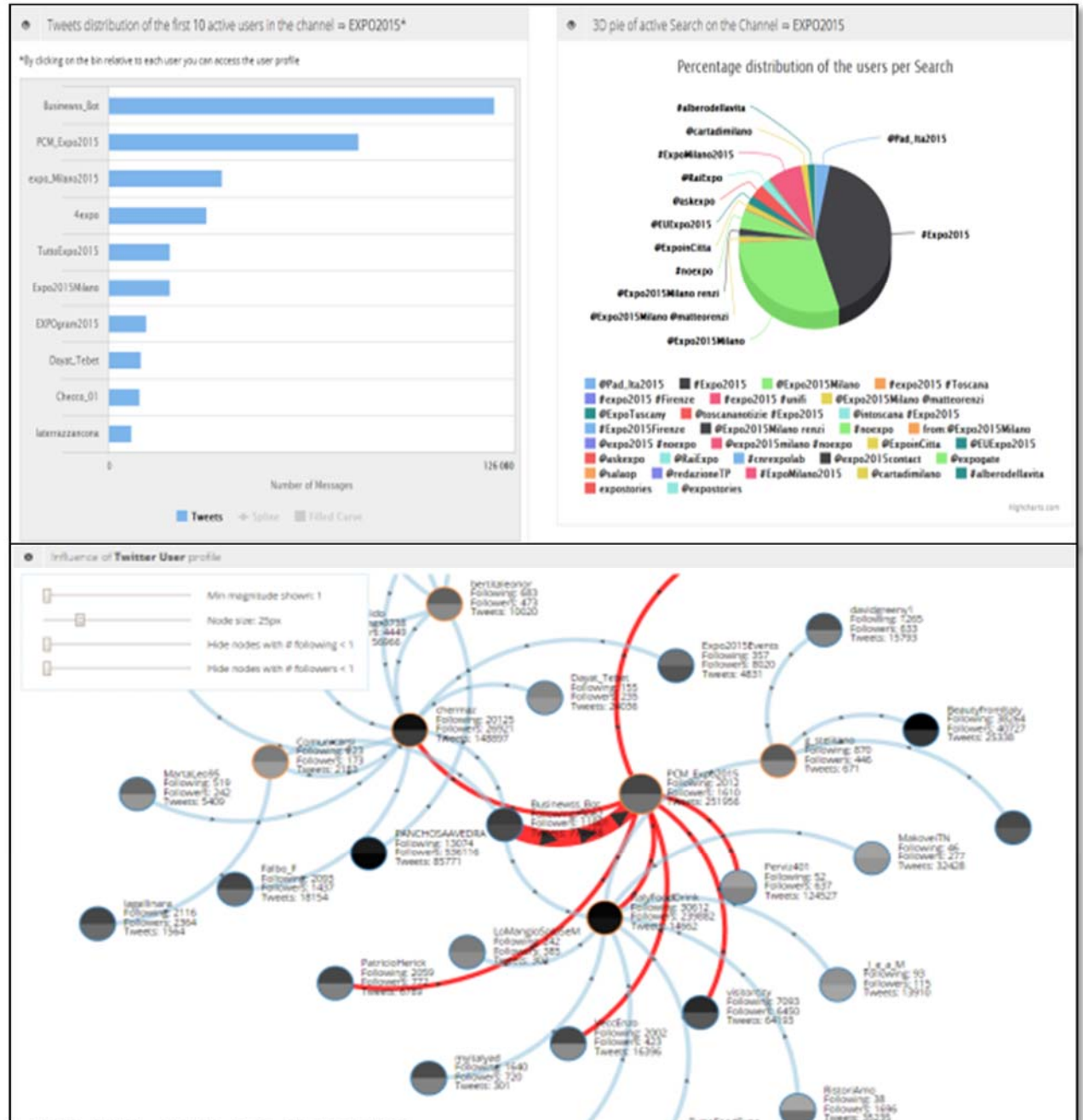
Un sentiment marcatamente negativo: -313272
Anche su key multiple

Pochi Tweet negativi hanno provocato un enorme numero di retweet

- **emergenza** caldo, una rete per i più deboli <http://t.co/x6e4dkwgk8>
- continua l'**emergenza** #caldo. domani fino a 43 gradi percepiti in #piemonte <http://t.co/gvuyafeq0o>
- cerolini su gestione **emergenza** caldo a pescara <http://t.co/zecmr7t4zv> #**emergenza**caldo #pescara
- **emergenza** caldo, in arrivo i 38 gradi nella città metropolitana di milano. consulta il bollettino emesso: <http://t.co/n5oocelibv>

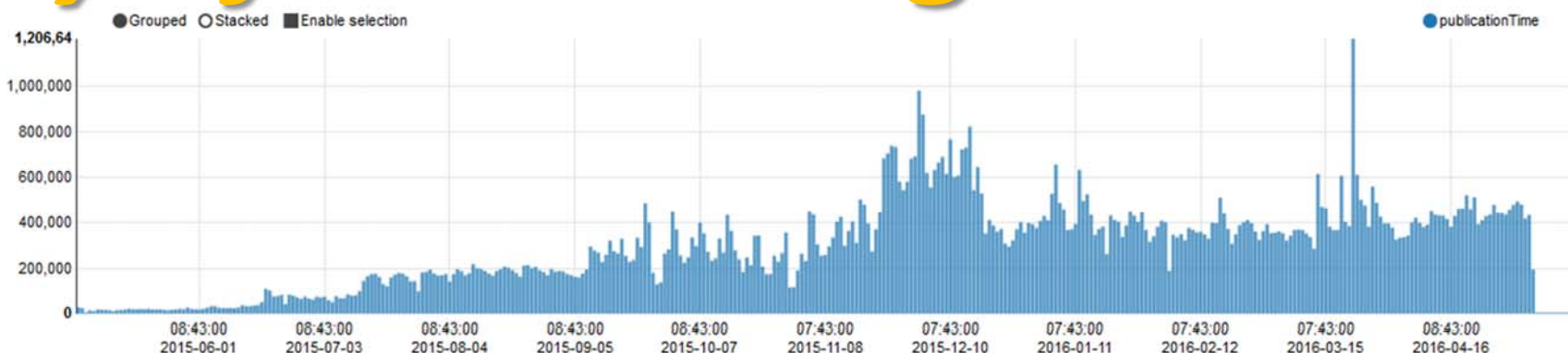


- Identification of
 - Influencer
 - Chains of users
 - communities
 - pusher
 - BOT
 - Relationships
- Assessing and predicting Virality

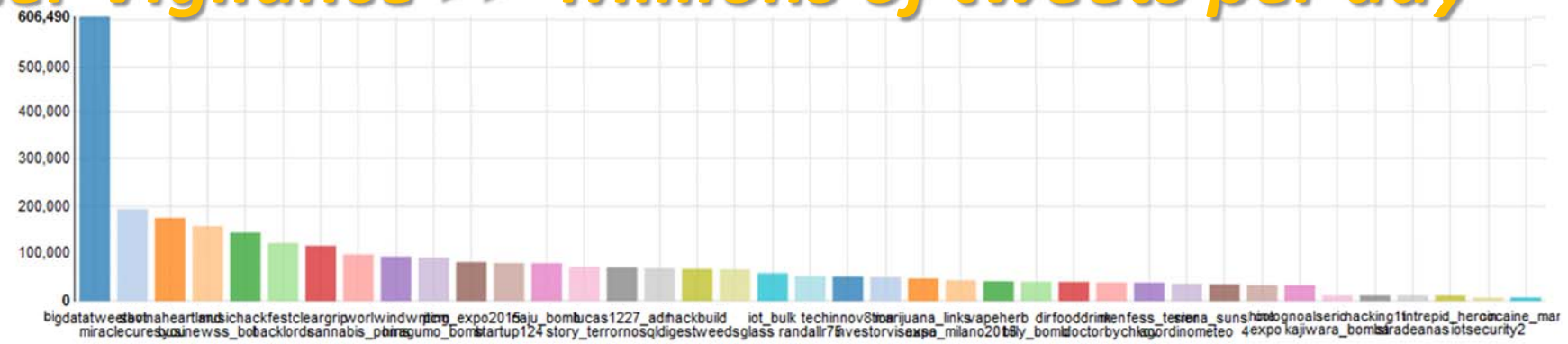


Full text and Faceted Search Facility of Twitter Vigilance db

- terrorismo (342208)
- farmacovigilanza (9740839)
- tech (8495537)
- uber (7337300)
- smart drugs (7062871)
- ads (91055)
- firer (9100)
- carlamer (198)
- relusso (303968)
- clima (2958966)
- xfactor9 (2672493)
- turismo firenze_2 (2451517)
- caldo (2395023)
- influenza (2374427)
- tpl (2051480)
- expo2015 (1643031)
- maltempo (1624816)
- allertameteo (1623214)
- cnrxpo (1541848)
- mymeteo (1275964)
- xylella (984595)
- allertameteo toscana (901664)
- turismo firenze (759910)
- turismo firenze_1 (759910)
- chi parla di meteo (699119)
- smartcity (646207)
- resilienza (636137)
- siena (496935)
- papafrancesco (485570)
- giubileo (413690)
- geocomunitas (399976)
- nasa new horizons (336838)
- laudatosi (321243)
- laudatosi_1 (321243)
- cardboard (294675)
- meteoalert (283981)
- nutella (252075)
- emergenza_acqua (245792)
- monitallergie (193804)
- mugnone2016 (175584)
- Show more...



Twitter Vigilance >> Millions of tweets per day



- twitterUser
- bigdatatweetbot (606490)
- miraclecuresyou (194035)
- savinahearthland (175355)
- businewss_bot (157586)
- musichackfest (144278)
- hacklords (122107)
- cleargrip (116207)
- snabbie_news (90956)

Grid Results

Filter fields

Tutti (65) / Current (2)

Field Name	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

[Http://tvsolr.disit.org](http://tvsolr.disit.org)

Twitter Vigilance



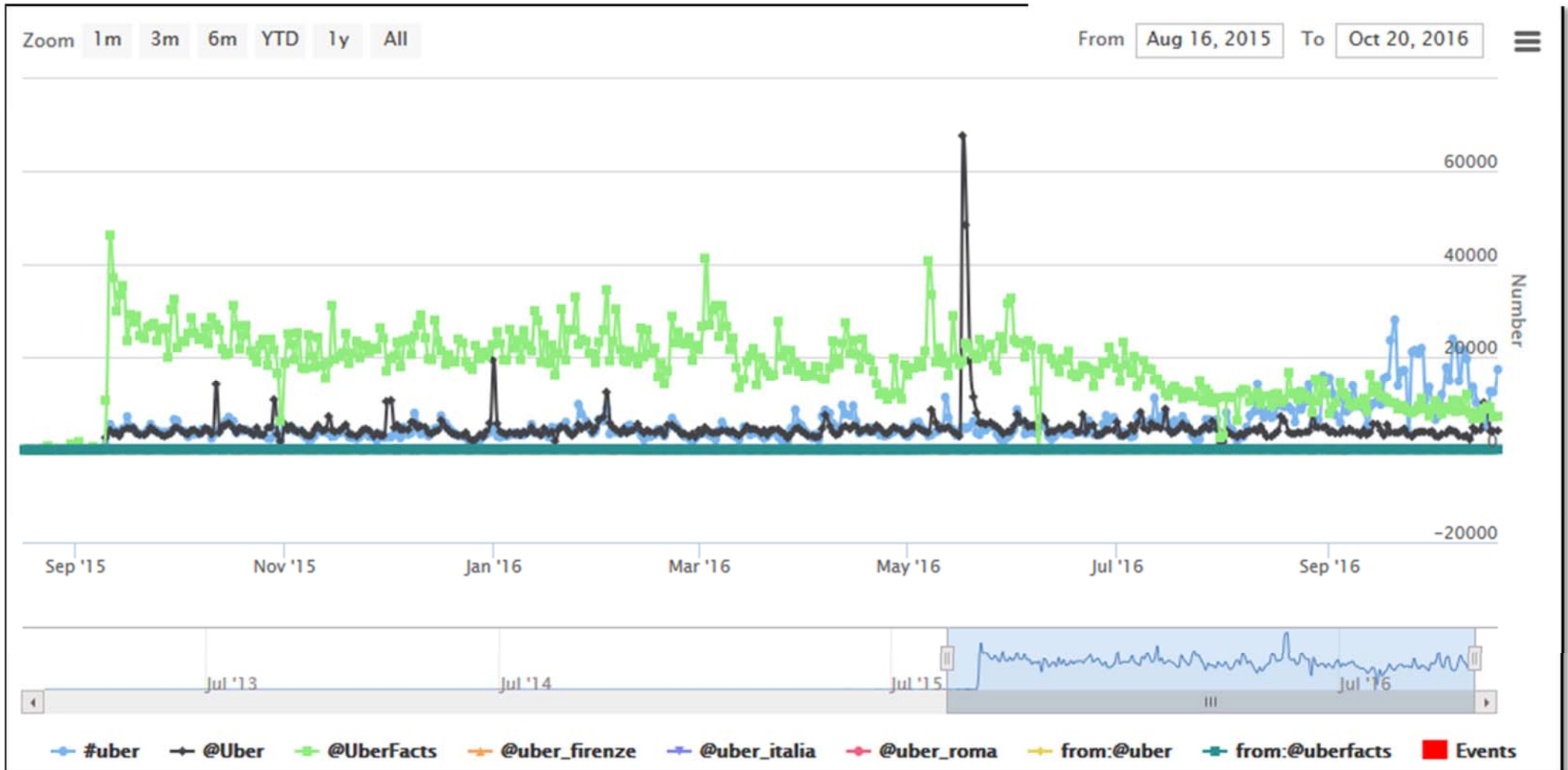
UNIVERSITÀ
DEGLI STUDI
FIRENZE

DINFO
DIPARTIMENTO DI
INGEGNERIA
DELL'INFORMAZIONE

DISIT
DISTRIBUTED SYSTEMS
AND INTERNET
TECHNOLOGIES LAB

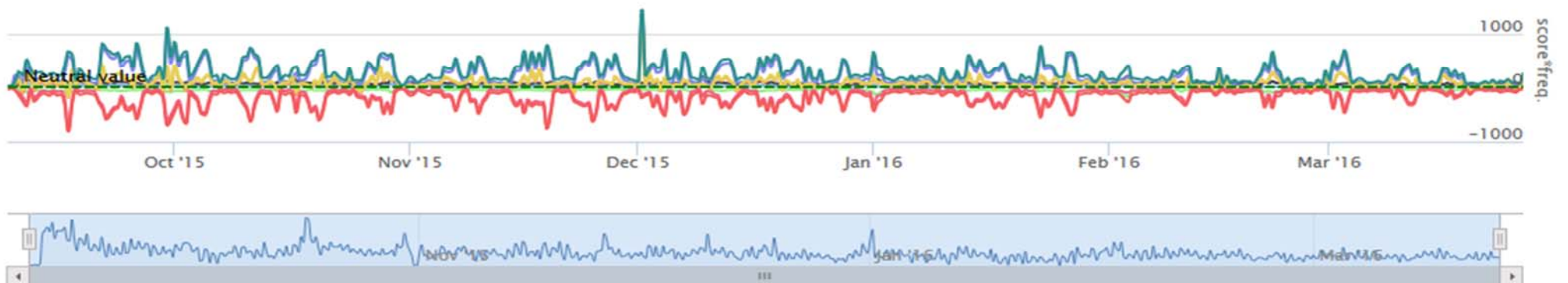
#uber | @Uber | @UberFacts
@uber_firenze | @uber_italia
@uber_roma | from:@uber
from:@uberfacts

Canale "Uber"



Zoom 1m 3m 6m YTD 1y All

From Sep 9, 2015 To Mar 28, 2016



— Tweets score
 — Tweets score pos
 — Tweets score neg
 — Retweets score
 — Retweets score pos
 — Retweets score neg
 — T+RT score
— T+RT score pos
 — T+RT score neg

Highcharts.com

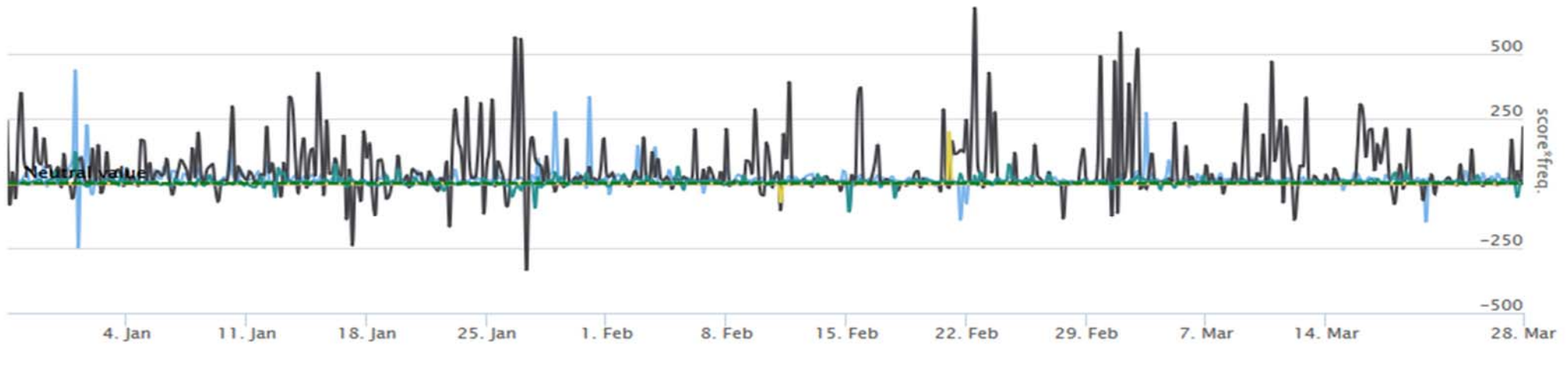
Hide All

Uber Sentiment Analysis

Sentiment trends in channel uber research

Zoom 1m **3m** 6m YTD 1y All

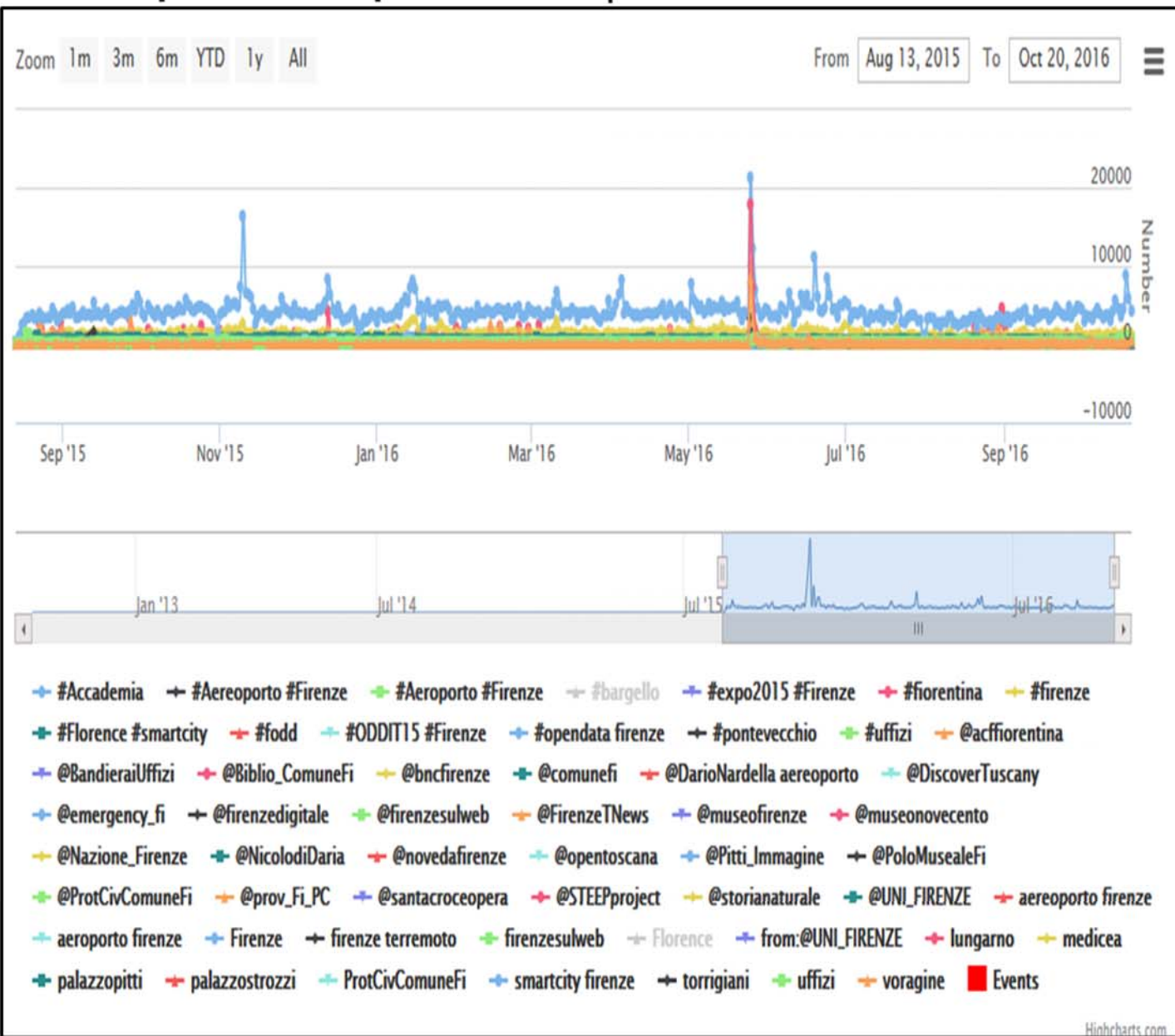
From Dec 28, 2015 To Mar 28, 2016



— @Uber
 — @UberFacts
 — @uber_roma
 — @uber_firenze
 — @uber_italia
 — from:@uber
 — from:@uberfacts
 — #uber



Canale "Firenze"

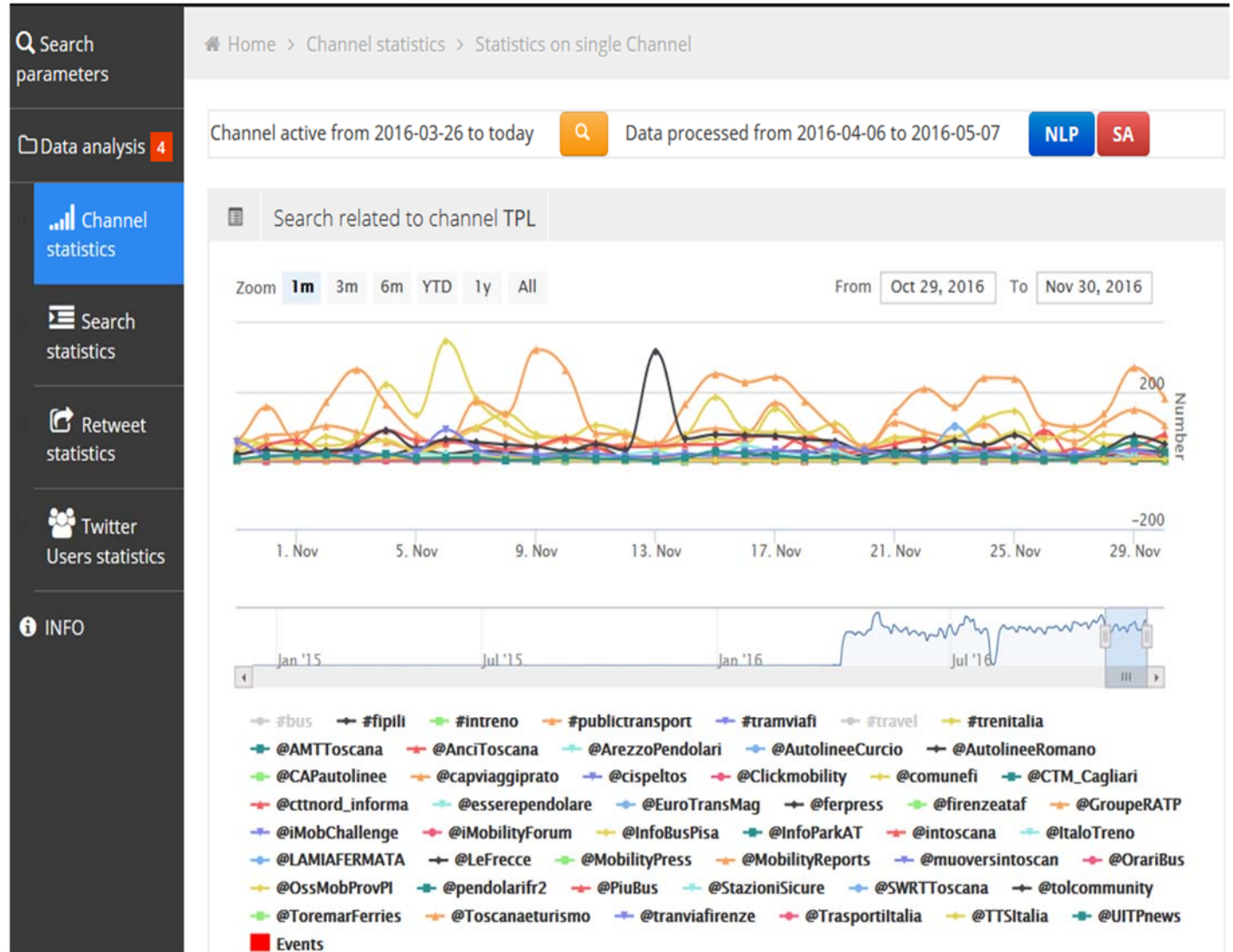


- #Accademia #Aeroporto #Firenze
- #Aeroporto #Firenze #bargello
- #expo2015 #Firenze #fiorentina #firenze
- #Florence #smartcity #fodd
- #ODDIT15 #Firenze #opendata firenze
- #pontevecchio #uffizi @acffiorentina
- @BandieraiUffizi @Biblio_ComuneFi
- @bncfirenze @comunefi
- @DarioNardella aeroporto
- @DiscoverTuscany @emergency_fi
- @firenzedigitale @frenziesulweb
- @FirenzeTNews @museofirenze
- @museonovecento @Nazione_Firenze
- @NicolodiDaria @novedafirenze
- @opentoscana @Pitti_Immagine
- @PoloMusealeFi @ProtCivComuneFi
- @prov_Fi_PC @santacroceopera
- @STEEPproject @storianaturale
- @UNI_FIRENZE aeroporto firenze
- aeroporto firenze Firenze
- firenze terremoto firenziesulweb Florence
- from:@UNI_FIRENZE lungarno medicea
- palazzopitti palazzostrozzi
- ProtCivComuneFi smartcity firenze
- torrigiani uffizi voragine



- Canale “TPL”, 12 M tweet

- #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
- @LeFrecece @MobilityPress
- @MobilityReports
- @muoversintoscana @OrariBus
- @OssMobProvPI @pendolarifr2
- @PiuBus @StazioniSicure
- @SWRTToscana @tolcommunity
- @ToremFerries @Toscanaeturismo
- @tranviafirenze @Trasportitalia
- @TTSItalia @UITPnews

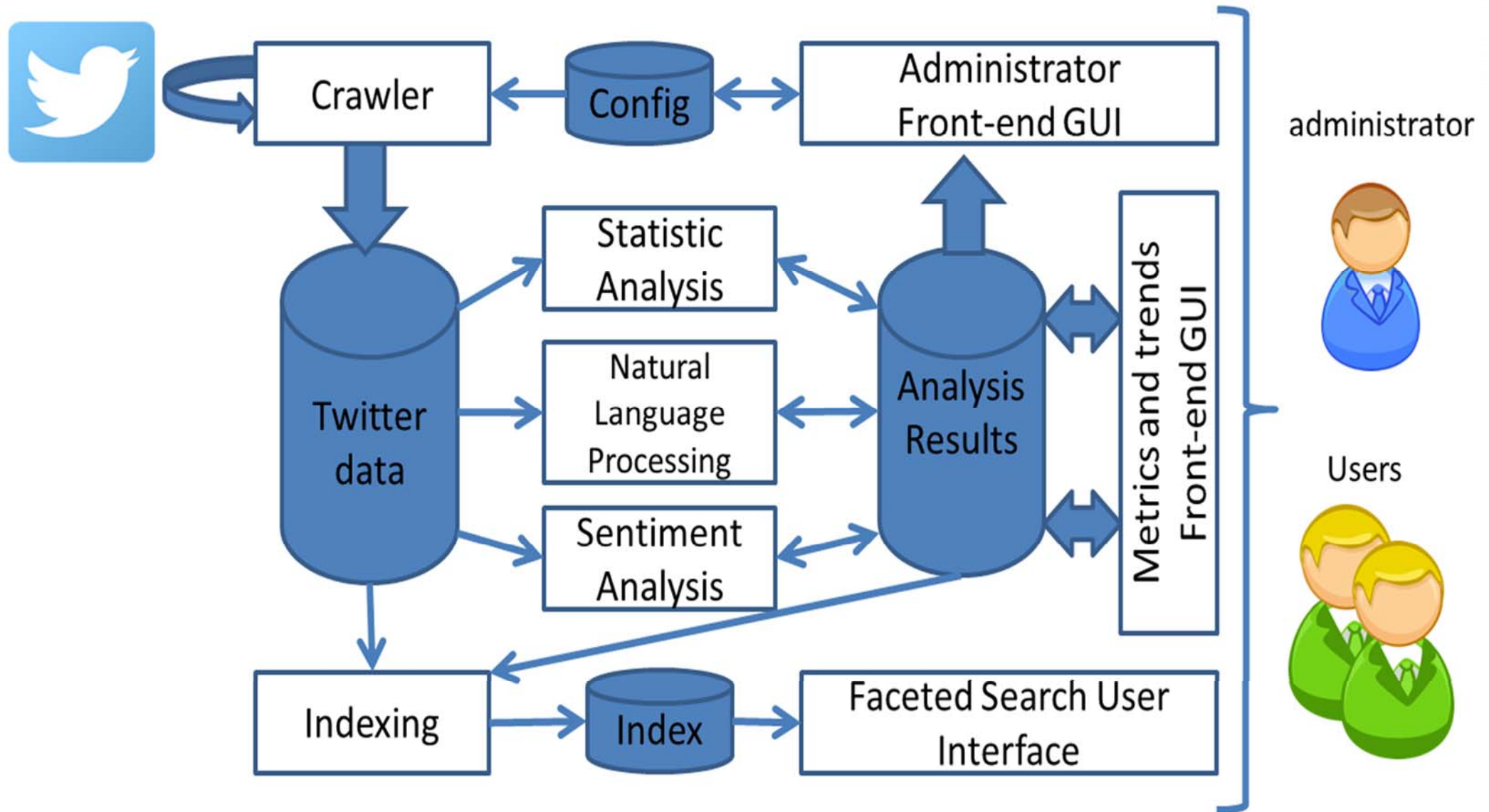




Twitter Vigilance: sentiment analysis

- **Controlla e analizza in automatico il *livello di apprezzamento e/o dissenso* per:**
 - prodotti, servizi, promozioni, cambiamenti in città,
 - persone, azioni politiche, eventi,
 - programmi TV, attori, cantanti (surrogato di "auditel")
- **Permette di effettuare**
 - Valutazioni di **andamento** a breve e lungo termine
 - Valutazioni **comparative** a breve e lungo termine
 - **Predizioni** in certe situazioni
 - **Identificazioni in quasi real time** della nascita di eventi esplosivi, situazioni critiche, etc., uso di utenti come sensori diffusi

Twitter Vigilance





Conclusioni

- famiglia di strumenti di Twitter Vigilance per l'automazione di analisi del canale social media Twitter
 - diagnosi precoce di problemi, sviluppo di modelli predittivi, studio dei social media, in vari domini, ecc.
- Molte sperimentazioni hanno permesso di crescere e diventare uno strumento efficace e semplice da utilizzare.
 - più di un anno di rilevamento e analisi.
- non sono la risposta definitiva all'analisi del social media o di dati Twitter, permettono di:
 - collezionare i dati in automatico,
 - produrre deduzioni semplici anche in automatico,
 - produrre indagini lasciando spazio per lo sviluppo di analisi dettagliate e approfondite dei dati collezionati.