



Erinn//

@sixpillowgeorge



Just a picture of me in bed last night..#TooHotToSleep

9:46 AM - 20 Jul 2016



**CARISMAND**

Culture And RiSk management in  
Man-made And Natural Disasters



Numero gratuito per  
informazioni sulle  
**Ondate di calore**

## Public crowd-sensing of heat-waves by social media data

*Heatwaves' impact in Italy during summer 2015 by  
twitter social media audit.*

**Valentina Grasso** <sup>(1,2)</sup>, **Alfonso Crisci** <sup>(1)</sup>, **Marco Morabito** <sup>(1)</sup>, **Paolo Nesi** <sup>(3)</sup> and **Gianni Pantaleo** <sup>(3)</sup>

(1) CNR Ibimet, Italian National Research Council, Florence, Italy  
(v.grasso@ibimet.cnr.it)

(2) LaMMA Consortium, Italian National Research Council, Florence, Italy.

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



16th EMS Annual Meeting & 11th European  
Conference on Applied Climatology (ECAC)  
| 12–16 September 2016 | Trieste, Italy

*Where atmosphere, sea and land meet:  
bridging between sciences, applications and  
stakeholders*

# This study was carried out in the field of the CARISMAND Project: *Culture And RiSkmanagement in Man-made And Natural Disasters*



**CARISMAND**  
Culture And RiSk management in  
Man-made And Natural Disasters

<http://www.carismand.eu/>

This project has received funding from  
the European Union's Horizon 2020  
research and innovation programme  
under grant agreement No 653748.



RESOURCES | NEWS AND VIEWS | EVENTS | CULTURAL MAP |



HOME ABOUT CARISMAND TEAM ACTIVITIES GALLERY MEDIA CENTRE CONTACTS

INTRANET

**LIVE** The earthquake in Italy: stereotyped narratives and missing social science

## About CARISMAND

- ✓ **Project Concept** - CARISMAND aims to deal with the issues of preparedness, response to disasters and after-crisis recovery which is inevitably influenced by cultural background...
- ✓ **Project Objectives** - Pursuing its goals towards culturally-informed solutions for disaster management...
- ✓ **Why is CARISMAND unique?** - CARISMAND comprehensively addresses a number of specific challenges and scope through a variety...
- ✓ **Who does CARISMAND concern?** - All CARISMAND activities are designed in such a way to be able to reach these

## News



### EMSC's 5 Visual Safety Tips for Good Practices after an Earthquake

You could help The Euro-Mediterranean Seismological Centre improve their safety tips by answering a few questions.



### AIDF Global Disaster Relief Summit, 7-8 September 2016

The summit is organised by the Aid & International Development Forum and will take place in Washington DC, USA.



### Deadly Italian Quake Highlights Continuing Struggle to Communicate Risk

An article by Edwin Cartlidge focusing on the issues of shaping a clear, non-misleading preparedness message to a population in the grip of a disaster.



**CARISMAND**  
Culture And RiSk management in  
Man-made And Natural Disasters



# Backgrounds

- Social media data become an unique **realtime informative source** to be used to monitor the impacts related of **meteoclimatic events**.
- **Heat wave** is the one of main issue in relation to the climatic hazard on **population**.
- The investigation on the reliability of **social media content** in relation to the crowd perception of **natural risk as heat-wave** is good topic.
- A **methodological approach** where data tools and framework give an effective exploitation of crowd-sensed data in regard to **heat related hazard** is a need today.



# Aims of work

- Are the environmental **hot conditions** ( heat-wave) associated to semantic related **social media (SM) streams**?
- Does **SM** activity “follow” the spatial and temporal patterns of **heatwaves** (HW) in Italy?
- Is a climatic daily classification ( as **HW days**) able to discriminate different level of local **SM activity**?
- Do exist a **social media** sensitivity linked to local **climatic parameters**?

These are the basic questions to define the reliability of social media analytics to perform a real-time monitoring the impact of heatwave.

*In the work will try to give some answers.*



# Weather based HeatWave status definition

It is a prolonged period of excessive heat, often combined with excessive humidity lasting 2-+ days. Critical day definition for a generic location and season is needed.



Image from:  
<https://www.er24.co.za/blog/heat-wave-be-cautious/>

- **EU-project EUROHEAT(\*)** approach was adopted improving local daily climatologies (1981-2010) . It states that Critical day occurs when:

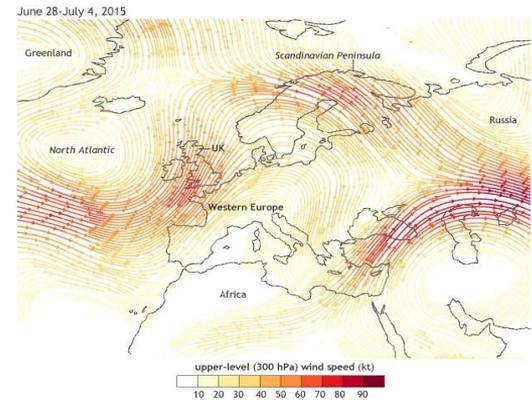
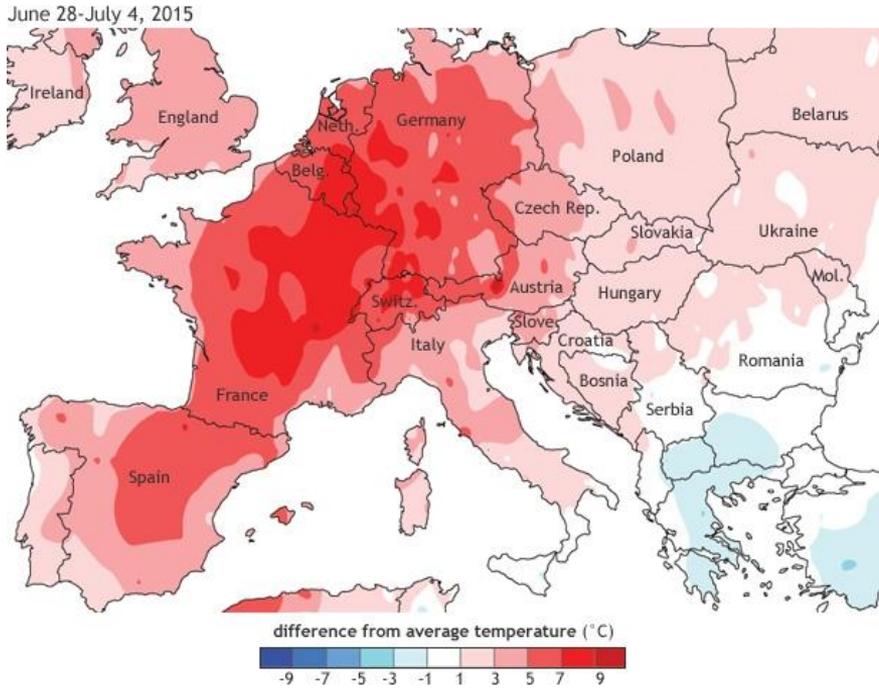
- $T_{\max \text{ Apparent}_{jd}} > T_{\max \text{ Apparent}_{jd} 90^{\text{th}}}$
- $T_{\min_{jd}} > T_{\min_{jd} 90^{\text{th}}} \ \& \ T_{\max \text{ Apparent}_{jd}} > \text{median } T_{\max \text{ Apparent}_{jd}}$

\* D'Ippoliti et(2010) Environ Health. "The impact of heat waves on mortality in 9 European cities: results from the EuroHEAT project."

- **Heat Wave** is when critical days spell exceed 2 days. The HW features are: duration (days) , intensity ( $T^{\circ} > \text{treshhold}$ ), seasonal precocity (jd timing) and spatial extent.

# Heat wave 2015 in Italy

Five main heatwaves episodes have occurred in Italy during summer 2015. The main one hits in July maintaining a persistent character for 1 months in reason to a atmospheric blocking ( "omega block"). North Thyrrenic regions are the most impacted in July. The agust's episode have involved the southern eastern areas ( mainly Apulia and Calabria).

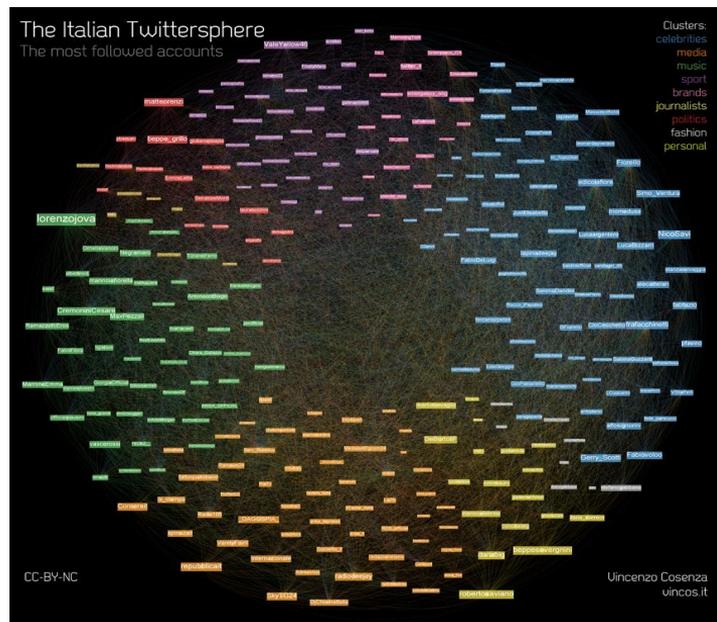


Web reference <https://www.climate.gov/news-features/event-tracker/summer-heat-wave-arrives-europe>

# Performing crowd-sensing by Twitter data

On-line multilanguage platform for social-networking and microblogging.  
Twitter data perform significant crowd-sensing.

**6.4 million** of active users in Italy (2015)



Who feed information in  ?

Many and various

Citizen, Institution, Institutional Public Services, Business Companies, Community - NGO organization, Media, Conversational Bots and Sensors Bots.



Daily activity common metrics

**RTW\_TW**: N° of tweets & retweets

**RTW**: N° of tweets & retweets

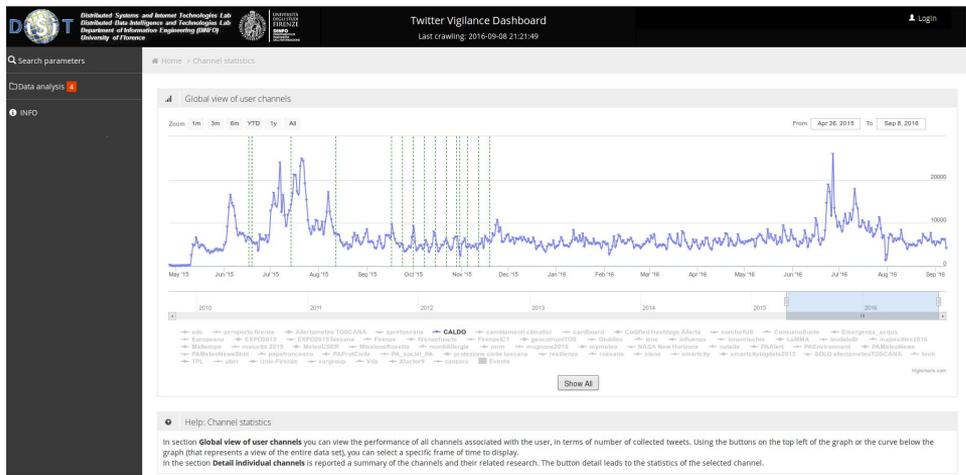
**TW**: N of native tweets & retweets

**U\_native\_users**: N° of native TW authors

**U\_full\_user**: N° authors of TW & RTW

**U\_unique\_hashtag**: N° of hash-TAG

# Twitter Vigilance DISIT Platform



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

Twitter Vigilance platform is an environment developed by DISIT University of Florence that:

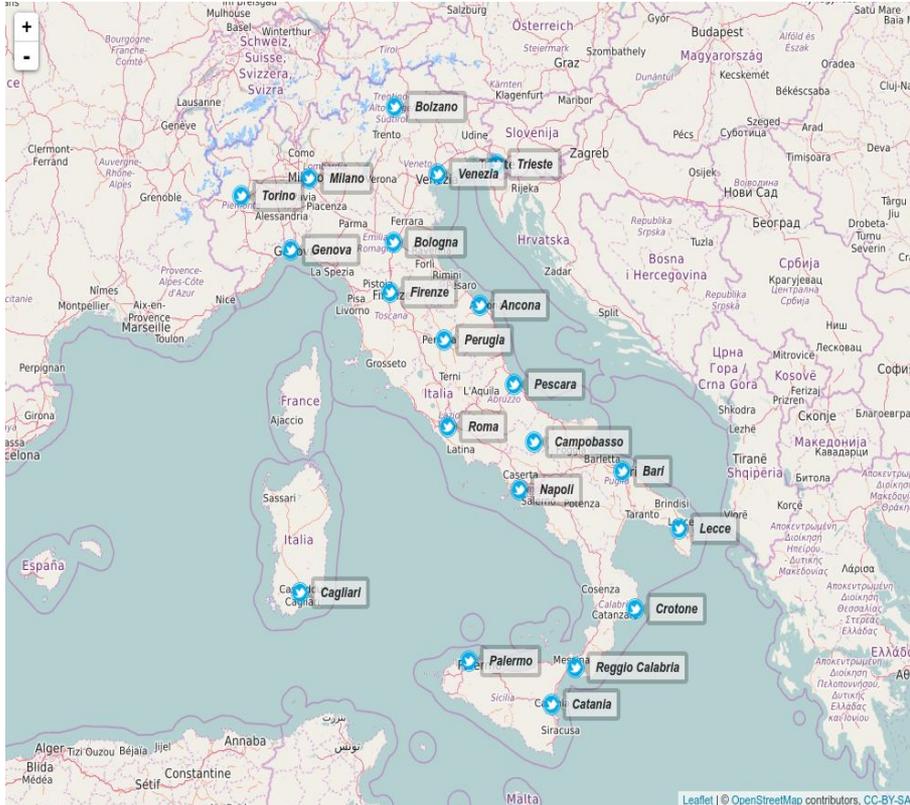
- Manage multiple queries in twitter API
- Store the data of messages collected by user defined queries -> channel
- It is a dashboard able to visualize data collecting process & analytics of twitter metrics of channel

Social media data: Channel “CALDO”

[http://www.disit.org/tv/index.php?p=chart\\_singlechannel&canale=CALDO](http://www.disit.org/tv/index.php?p=chart_singlechannel&canale=CALDO)

Channel	Related research	Total	N° tweets	N° tweets(%)	N° retweets	N° retweets(%)	Details	Analysis
CALDO	#caldo #sole afa allerta allerta caldo anticiclone caldissimo caldo canicola disagio caldo emergenza caldo estate rovente flegetonte ondata di calore rischio rovente temperature torrido	3344111	1851650	55.37%	1492461	44.63%	From 2009-10-23 To today	From 2015-05-15 To 2015-09-15

## Location considered : workflow weather for twitter data.



- GSOD [1] daily weather data for 21 italian location are acquired corresponding to main city for each region.
- Daily Max  $T^{\circ}$  Apparent are assessed by using Steadman [2] approach.
- Daily normals & quantile (1981-2010) are performed for Tmin and Tmax by using 31-day window for JD.
- Critical days and Heatwave days are calculated for the period ( 2015-05-15 / 2015-05-15). N=124.
- Statistics for each region are calculated.

**Climatic analysis are carried out to link outcomes with social media data for the period investigated (15 May to 15 September). Main city are considered as representative of own region.**

[1] <https://data.noaa.gov/dataset/global-surface-summary-of-the-day-gsod>

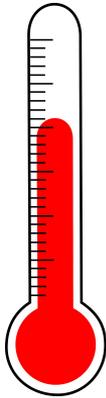
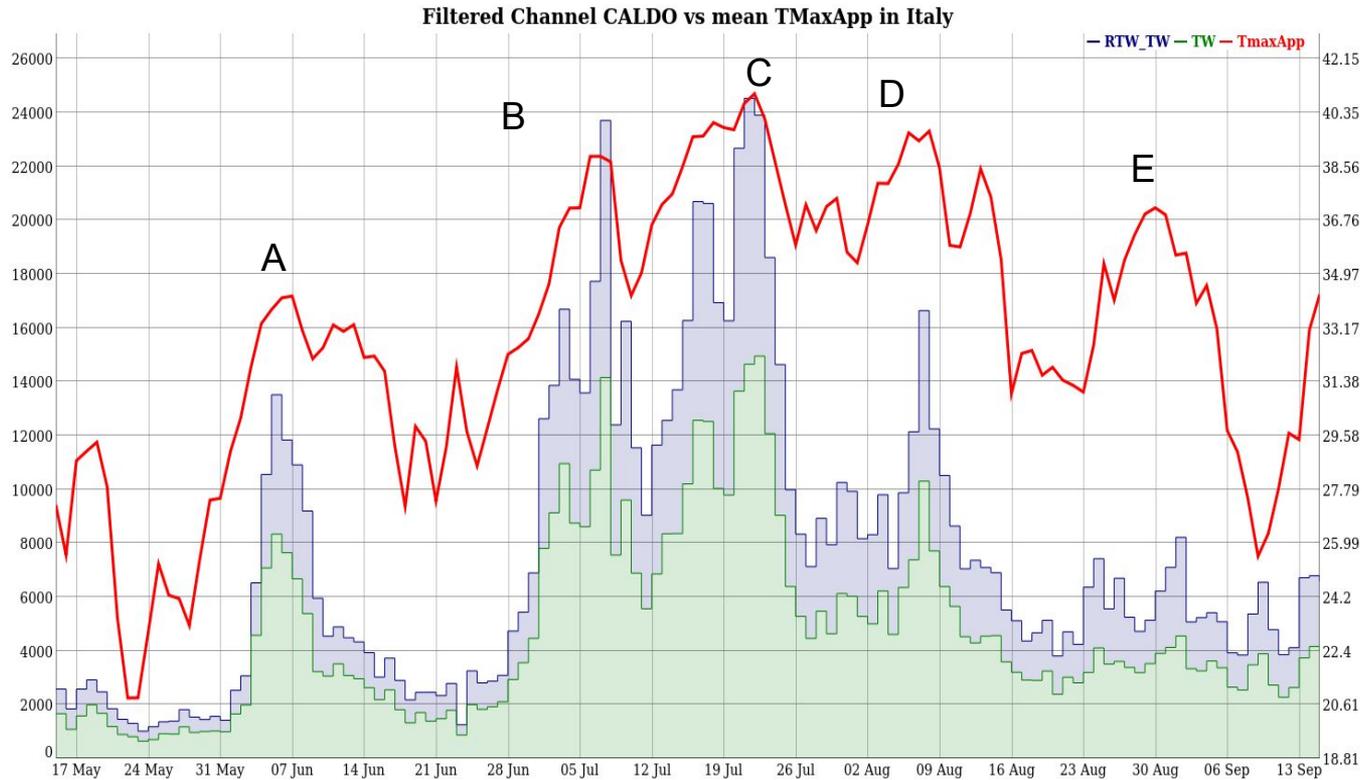
[2] Steadman, R.G. 1984. A universal expression of apparent temperature. J. Appl. Meteor. **23**, 1674-87

# Channel “caldo” data and geo-extraction of local twitter streams.

Area	RTW_TW	TW	RTW	retweet	N_users	F_users	N_hashtag	Most_retweeted
<i>marche</i>	106	77	29	152	51	73	55	ITnewsAN
<i>puglia</i>	934	473	461	25581	306	716	256	99Ludo MNW_Puglia LucaLombroso
<i>emilia</i>	606	461	145	850	313	433	220	alto_adige
<i>trentino</i>	221	171	50	162	121	157	93	paoloigna1
<i>sardegna</i>	648	457	191	1279	295	452	195	mangiareinmolis
<i>molise</i>	61	53	8	28	29	37	24	paynehoran16
<i>calabria</i>	916	228	688	507633	180	859	137	flash_meteo
<i>toscana</i>	1662	1098	564	3133	487	760	397	Ossmeteobargone
<i>liguria</i>	1320	910	410	1896	303	519	314	besidenjall
<i>lombardia</i>	2944	1949	995	50893	1391	2152	1075	mariamdj202
<i>sicilia</i>	1785	874	911	222818	622	1449	412	ItnewsPG
<i>umbria</i>	353	247	106	495	96	144	90	DonatellaPetre1
<i>abruzzo</i>	240	181	59	324	98	138	111	TrastevereRM
<i>lazio</i>	2705	1834	871	11851	1308	2060	868	ArpaPiemonte
<i>piemonte</i>	1145	698	447	3480	418	758	313	chermaz
<i>friuli</i>	218	134	84	411	85	140	97	Gazzettino
<i>veneto</i>	589	463	126	577	283	392	224	
Totals	16453	10308	6145	831563	6386	11239	4881	
Italia	940123	585286	354837	49113594	159917	233553	61385	VoiceItalia

- Original Twitter Channel “Caldo” are filtered matching message containing the heat related words ["caldo", "afa", "canicola", "sudore", "caldissimo", "torrido", "record", "allarme", "emergenza", "bollino", "bere", "anziani", "sete", "umidità", "sudore", "anticiclone", "disagio", "canicola", "caronte", "umido", "flegetonte", "suda", "sudato"] . Filtered Channel “Caldo” is indicated in table as “Italia”.
- Local streams are originated by a further geographical key-term filtering by the using terms linked to location considered [“city name”, “region name”].
- For each tweets collection the daily metrics of activity is calculated.
- Daily data are merged with the respective local weather data.

# Channel "Caldo Italia" and Heatwave periods 2015 (15 May to 15 September)



# The long and intense HeatWaves 2015: a very deep impact in Italy.

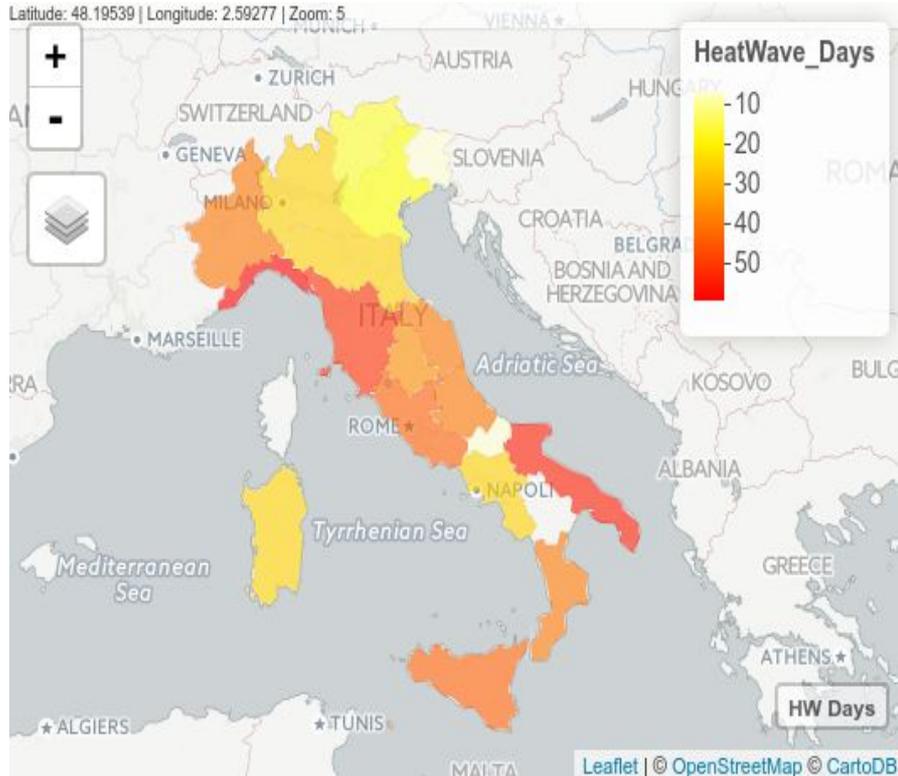
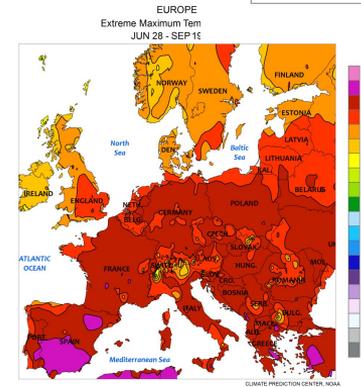
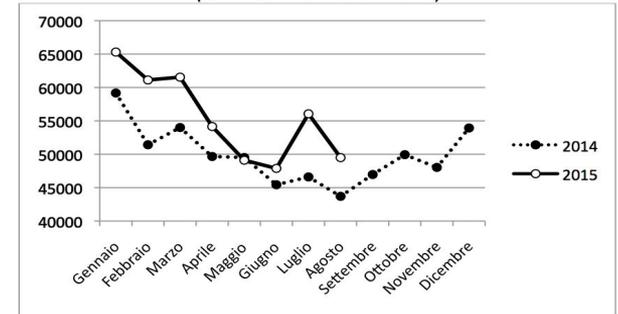


Fig. 1 - Italia: Frequenza mensile di morti. Anni 2014 e 2015  
(Fonte [www.demo.istat.it](http://www.demo.istat.it))



<http://www.neodemos.info/68-mila-morti-in-piu-nel-2015/>

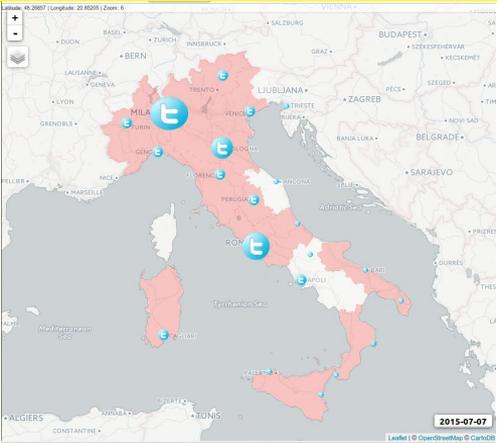
Image from Wiki Commons 2015 European Heat wave.

[https://commons.wikimedia.org/wiki/Category:2015\\_european\\_heat\\_wave?uselang=de](https://commons.wikimedia.org/wiki/Category:2015_european_heat_wave?uselang=de)

# Social Response Pattern during peak days

*Twitter social response are influenced not only by heatwave conditions but it is also strongly dependent by population density.*

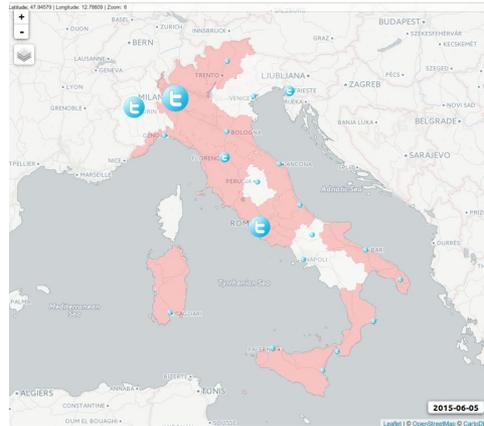
B July



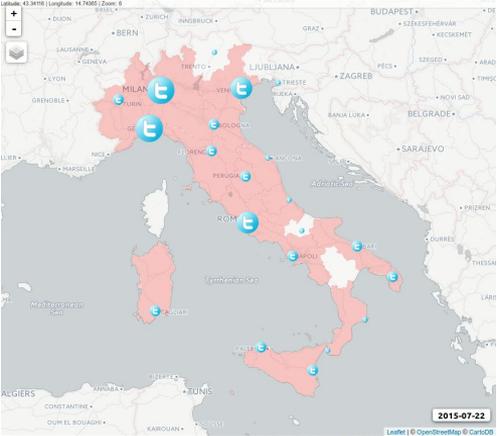
D August



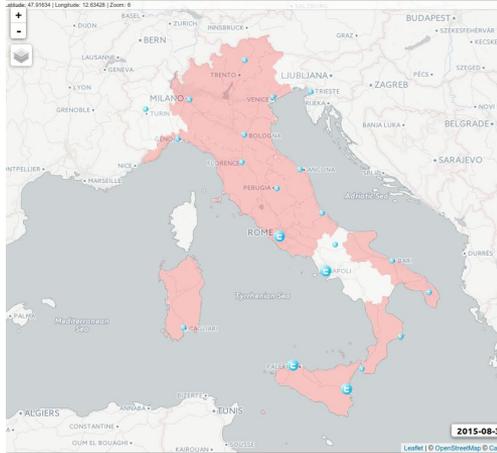
A Jun



C July

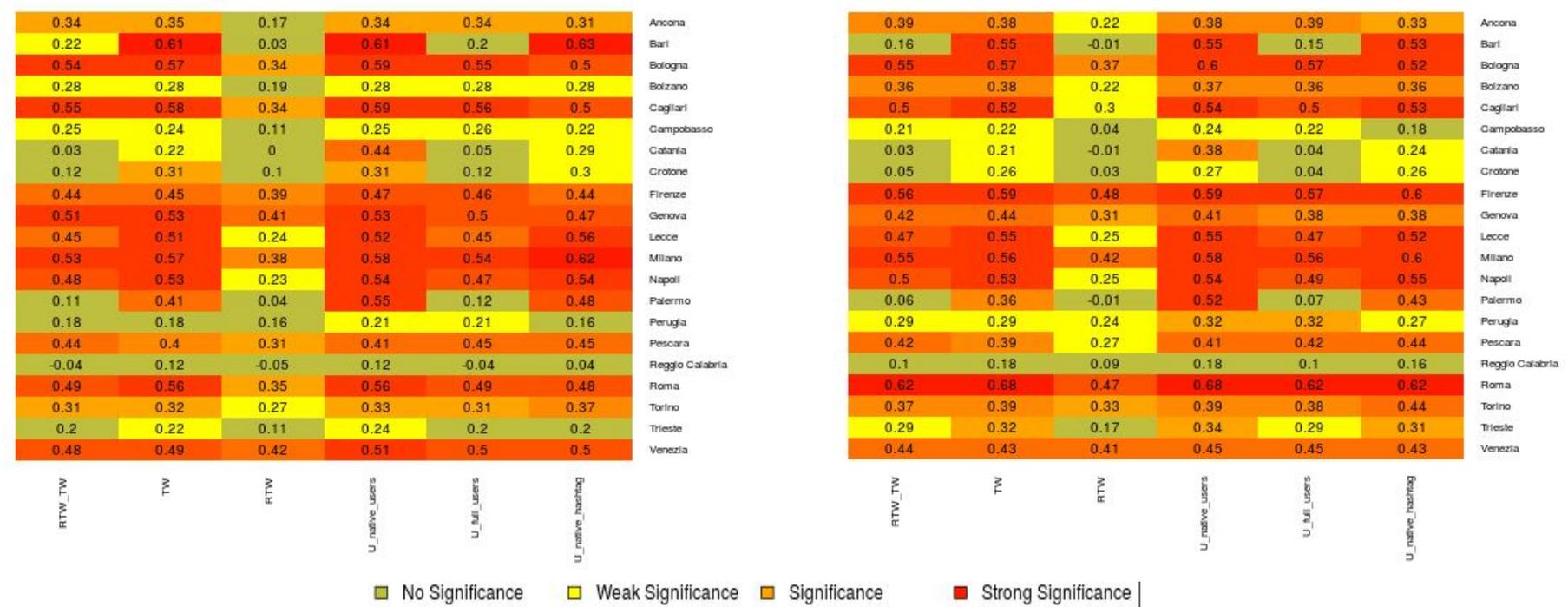


E August



# Association between daily Twitter Metrics ( RTW\_TW, TW, RTW, native\_users, full\_users, N hashtag) and main HW parameters by cities.

Linear Correlation between daily values in 15 may-15 september. N=124 . Minimum Level of significance  $r(0.05) \sim 0.15$

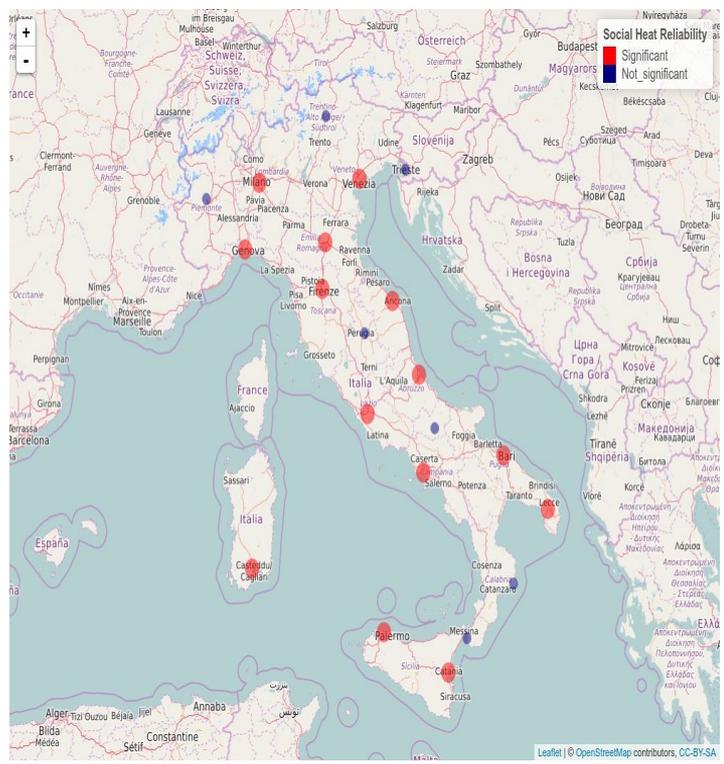


Daily Minimum temperature

Daily apparent maximum temperature

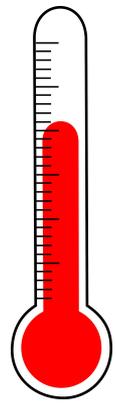
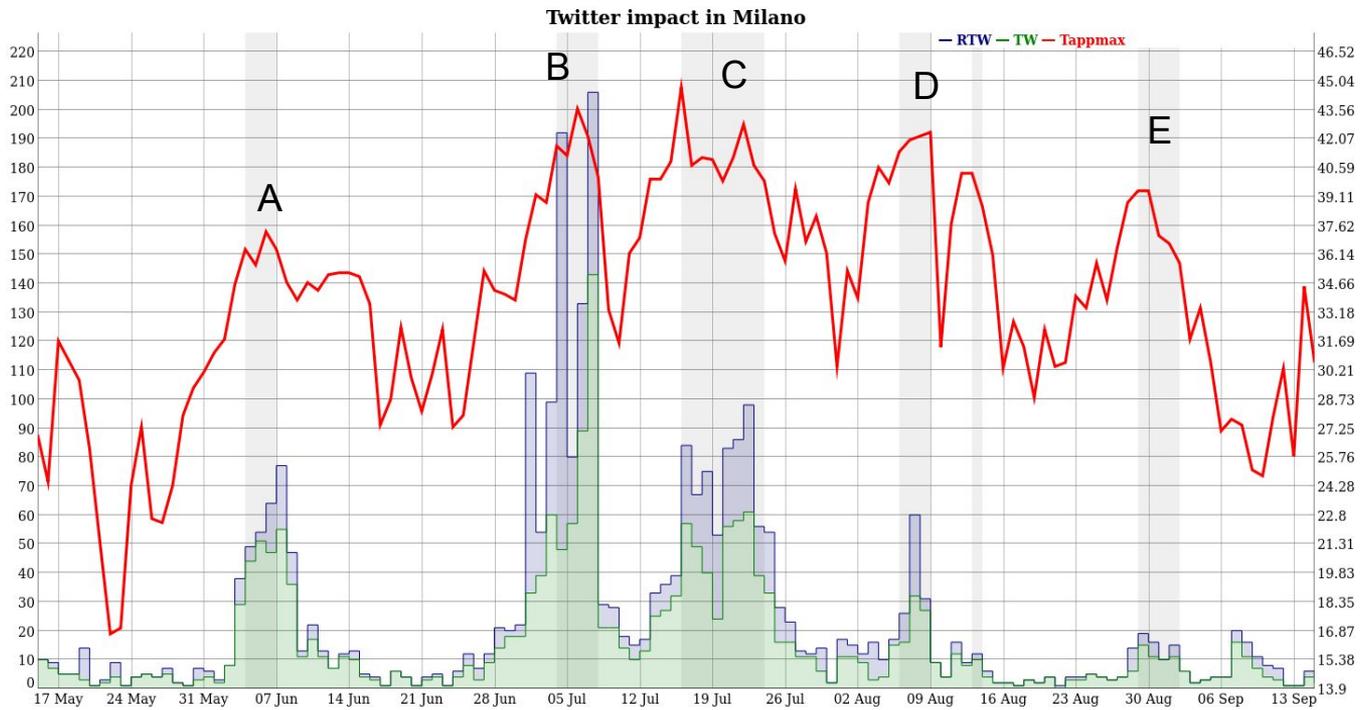
# HeatWaves and daily social reliability (how & where: t-student comparison & map)

City (*Significant)	All tweets		Native tweets		Native Users		Population
	NO HW	HeatWave	NO HW	HeatWave	NO HW	HeatWave	
Ancona *	0.5	1.6	0.4	1.2	0.4	1.2	100861
Bari *	5.1	10.7	2	6	1.9	5.6	1263820
Bologna *	2.9	10.9	2.1	8.6	2	7.8	1005831
Bolzano	1.6	1.9	1.2	1.6	1.1	1.5	106441
Cagliari *	3.3	11.4	2.3	8	2.2	7.6	154460
Campobasso	0.5	0.8	0.4	0.8	0.4	0.8	225622
Catania (*)	13.9	10	5.2	6.7	4.6	6.6 *	1115535
Crotone	9.7	4.8	1.6	2.2	1.5	2.1	62178
Firenze	8.5	20.8	5.7	13.6	4.7	11.5	1013348
Genova *	6.5	14.8	4.6	10.1	3.7	7.4	854099
Lecce *	1.8	5.8	1.4	3.7	1.3	3.6	94773
Milano *	11.9	60.3	8.2	38.7	7.8	35.3	3208509
Napoli *	5.8	14.3	4.5	11	3.9	9	3113898
Palermo *	14.2	14.8	5.7	9.5	5	9.2	1271406
Perugia	2.6	3.5	1.8	2.6	1.5	2.2	662110
Pescara *	1.1	3.8	0.9	2.8	0.8	2.4	121014
Reggio Calabria	8.7	4.3	1.5	2.1	1.4	2	555836
Roma*	11.5	41.3	8.2	27.3	7.5	25.1	4340474
Torino	7.8	13.2	4.9	7.6	4.6	7.1	2282197
Trieste	1.8	1.4	1.1	1	1	1	234874
Venezia *	2.7	14.5	2.2	11.1	2.1	9.9	855696

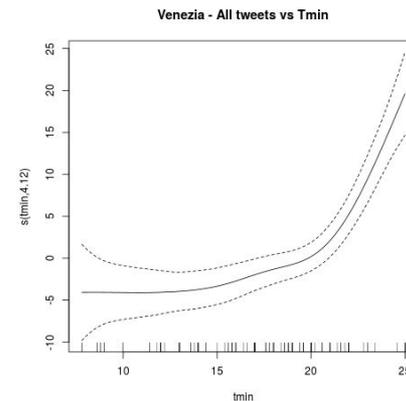
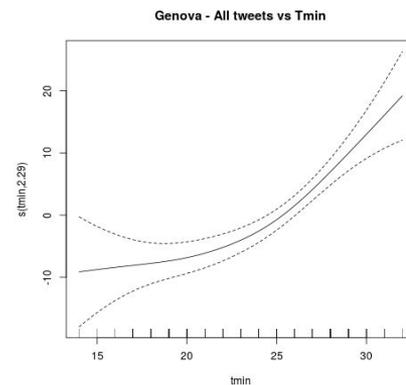
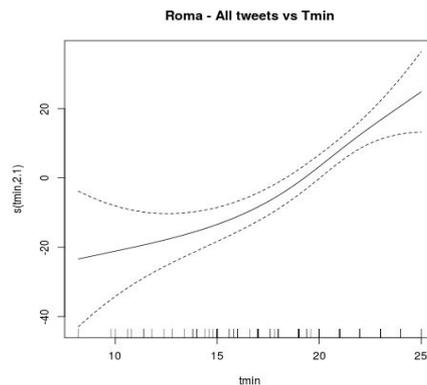
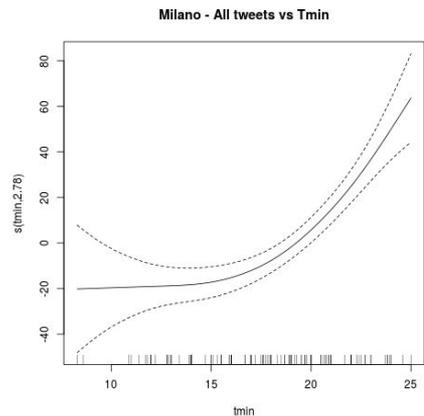
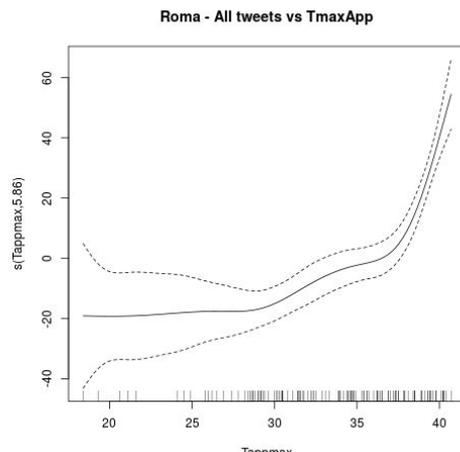
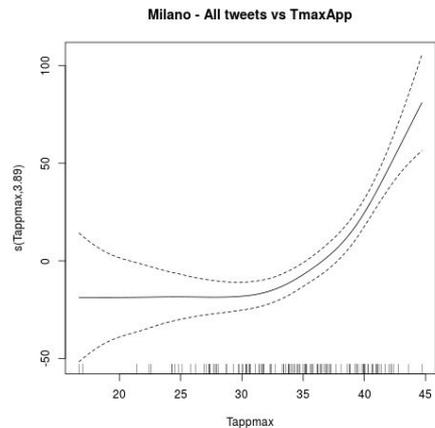


Leaflet | © OpenStreetMap contributors, CC-BY-SA

# HeatWaves and daily twitter activity Milano



# Parametric sensitivity of local daily twitter flows (GAM models Threshold is X the when $Y \geq 0$ )



Inlands cities show threshold around  $35^{\circ}\text{C}$  in TMax App.

Analysis suggest that coastal areas have more higher threshold in Tmin ( $25^{\circ}\text{C}$  and more).

These considerations are valid where association between HeatWave and tweets is significant. Further investigations are needed.

## Seven points remarks

- **Social media (SM)** data collection require right search-terms to obtain suitable data. the semantic tuning of **TW channel** is ever required. Platforms as **TwitterVigilance** are useful.
- Channel as “**Caldo**” have followed well the italian summer thermal profile in 2015. **HW events** are easily recognizable following **channel activity**.
- Simple local **SM stream** obtained by geo-term filtering, could be reliable only in a medium-large city context ( > 800.000 ab) to reach an adequate **SM data** numerosity.
- **SM data** suffers a strong dependence with population density. This is a considerable limit when **SM data** are used outside of **big metropolitan areas**.
- When data numerosity is adequate the association **SM activity** vs **HW parameters** is detected as significant. Moreover **HW status** ( with some exception) implies ever higher level of **SM activity**.
- **SM flows** exhibits often a spatial pattern following the timing and the extent of **HW phenomenon**.
- **SM sensitivity** activity vs **HW parameters** is statistically detectable. In **large urban areas** **SM data** could be an help to evaluate quantitatively the impact of phenomenon as **HW and disaster response**.

