

TORCIA: una piattaforma collaborativa per la gestione delle emergenze

Chiara Francalanci - Paolo Giacomazzi

1. Introduzione

TORCIA è un progetto finanziato da Regione Lombardia per lo sviluppo di una piattaforma che valorizzi l'utilizzo dei social media nei processi di gestione delle emergenze e, più in generale, nei processi di pianificazione e controllo delle situazioni critiche - incidenti o calamità naturali - garantendo la resilienza delle infrastrutture di trasporto sia per quanto riguarda l'infrastruttura di telecomunicazione sia quella della rete stradale. L'elaborazione del progetto ha avuto inizio nel giugno 2012 e si è conclusa nel giugno 2014. La partnership del progetto è composta da Alcatel-Lucent, capofila di TORCIA, dal Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB) del Politecnico di Milano, che ha la responsabilità scientifica del progetto, da Fondazione Politecnico di Milano, e dalle società ACT Solutions, Beta 80 e Vidiemme.

I Comuni italiani, sulla base delle proprie realtà e criticità territoriali, pianificano la risposta alle possibili emergenze tramite i piani comunali di protezione civile, che prevedono una serie di operazioni, ruoli e misure da adottare al verificarsi di una determinata calamità naturale o antropica. La viabilità è uno dei punti chiave da pianificare in caso di emergenza: il piano comunale di protezione civile infatti può prevedere la modifica del senso di circolazione e lo sgombero di determinate arterie per consentirne l'utilizzo ai soli mezzi di soccorso. Mantenere le condizioni di sicurezza e indicare la percorribilità dell'infrastruttura stradale sono elementi chiave per una corretta gestione delle situazioni di emergenza.

In questo contesto, il progetto TORCIA mira a facilitare le operazioni di raccolta dei dati sul territorio per consentire un'immediata comprensione delle situazioni critiche e delle iniziative da compiere. I cittadini utenti del Web 2.0 diventano protagonisti nel fornire, tramite i social network - facebook e twitter in primis - informazioni georeferenziate e dati utili per una migliore gestione della criticità in atto. Per raccogliere, razionalizzare e interpretare queste informazioni e condividerle con le autorità competenti, il progetto TORCIA ha messo a punto una infrastruttura di elaborazione basata su piattaforme cloud geograficamente distribuite e collegate tra loro da una rete ottica ad alta capacità. In questo modo TORCIA è in grado di sfruttare la collaborazione della cittadinanza analizzando i dati condivisi, utilizzando in modo tecnologicamente innovativo, integrato e completo la rete dei social network, sia per interpretare le informazioni in fase di emergenza, sia per fornirle alle autorità. La Figura 1 mostra l'architettura tecnologica complessiva del progetto Torcia.

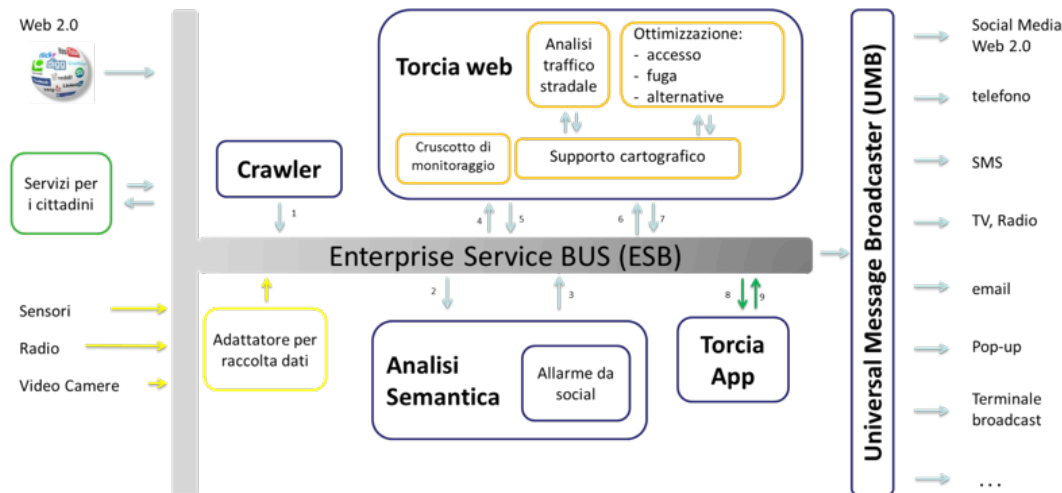


Figura 1
L'architettura della piattaforma TORCIA

2. L'informazione sociale

Il workflow di TORCIA parte con la raccolta continua e in tempo reale di informazioni dai principali social media. La prima domanda che occorre porsi è se i social media offrano effettivamente informazione pertinente al contesto del quale ci si sta occupando e utile per il raggiungimento degli obiettivi che ci si è posti. Nel nostro caso, è possibile rispondere positivamente a questa domanda se gli utenti dei social media, ovvero i cittadini, parlano di situazioni di emergenza sui social media e se tale informazione può essere effettivamente utile ai fini della gestione dell'emergenza stessa. La verifica di queste condizioni è stata oggetto di una interessante analisi preliminare di fattibilità che è stata condotta all'inizio del progetto.

La prima, ovvia, considerazione da fare è che se un'emergenza si verifica in un luogo scarsamente popolato, è molto probabile che non ci sia un numero sufficiente di utenti social media fra le persone coinvolte e che, di conseguenza, i social media siano poco informativi. Ad esempio, una frana di un sentiero a bassa percorrenza riceverà poca o nessuna attenzione sui social media da parte delle pochissime persone coinvolte. Viceversa, una frana di un sentiero famoso (ad esempio la frana della Via dell'Amore alle Cinque Terre) sarà tempestivamente segnalata sui social media. Generalizzando, i social media diventano tanto più utili quanto maggiore è il numero di persone coinvolte. Ci siamo dunque concentrati sul caso più favorevole, quello delle *emergenze cittadine* e, in particolare, delle *alluvioni*, divenute tristemente frequenti e perciò adatte a fornire numerosi test case in corso di progetto.

Con riferimento alle alluvioni in un contesto cittadino, le nostre analisi preliminari si sono poste i seguenti obiettivi:

- verificare i volumi di parlato su Twitter relativo alle alluvioni;
- verificare gli argomenti del parlato su Twitter relativo alle alluvioni;
- verificare la presenza di elementi utili alla geolocalizzazione del parlato in caso di alluvioni;
- verificare la presenza e il ruolo delle istituzioni su Twitter;
- confrontare la situazione italiana con le best practice internazionali.

ALLAGAMENTO	7	INONDATA	41
ALLAGAMENTI	58	INONDANDO	1
GRANDINE	74	ALLAGARE	29
GRANDINATA	12	ALLAGATO	40
GRANDINATE	2	ALLAGATA	79
TEMPESTA	337	ALLAGANDO	16
ALLUVIONE	30	ESONDARE	1
INONDAZIONE	3	ESONDATO	1
INONDAZIONI	31	ESONDANDO	0
ESONDAZIONE	1	FRANARE	7
FRANA	721	FRANATO	2
FRANE	52	FRANATA	10
URAGANO	69	FRANANDO	0
NUBIFRAGI	15	ALLUVIONATO	0
NUBIFRAGIO	47	ALLUVIONATE	0
INONDARE	4	ALLUVIONATI	6
INONDATO	20	ALLUVIONATA	1
		GRANDINE	8

Figura 2

Parole chiave per la raccolta di informazioni da Twitter (crawling keyword).

Il crawling, ovvero la raccolta di post, è stato attivato nel settembre 2012 su circa 60 keyword che definiscono il dominio alluvioni (in Italiano e in Inglese, fra cui allagamento, frana, alluvione, tempesta, uragano, inondazione, nubifragio, hurricane, flood, eccetera). La lista delle keyword in Italiano è riportata a titolo esemplificativo nella Figura 2. Complessivamente, sono stati analizzati oltre 40 milioni di tweet nel corso del progetto. I volumi riportati nella Figura 2 rappresentano il numero di occorrenze in una giornata di maltempo tipica (priva di situazioni che si possano definire di emergenza). Si nota come il parlato tenda a concentrarsi su alcune parole chiave, che, tuttavia, tendono a cambiare nel tempo a seconda sia delle condizioni di maltempo, sia dell'effetto di imitazione nella scelta delle parole che si crea nel momento in cui gli utenti si scambiano opinioni sul maltempo. Si può concludere che il parlato in lingua italiana sul dominio delle alluvioni ha volumi considerevoli (in media 40.000 tweet/mese) e può essere oggetto di ulteriori analisi (vedi anche [4]). Il volume di parlato in lingua Inglese è di un ordine di grandezza superiore [5].

Abbiamo poi ricavato *bottom-up* dall'analisi del parlato gli argomenti di conversazione più frequenti [3]. Il nostro obiettivo in questo caso era capire se il parlato può offrire indicazioni potenzialmente interessanti nella gestione delle emergenze. Le categorie di parlato più frequenti si sono rivelate le seguenti:

- *Dove*, ovvero indicazione di luoghi nei quali si sta verificando una situazione di maltempo.
- *Segnalazione*, ovvero indicazione di situazioni di maltempo e valutazione della loro gravità (da semplice disagio a vero e proprio allarme).
- *Conseguenze*, ovvero indicazione di fatti rilevanti accaduti come conseguenza del maltempo (allagamenti, crolli, frane, eccetera).
- *Responsabilità*, ovvero opinioni sull'attribuzione della responsabilità delle conseguenze del maltempo.
- *Viabilità*, ovvero conseguenze del maltempo sul traffico e sulla viabilità.
- *Warning*, ovvero discussione preliminare al maltempo di previsioni meteorologiche negative.
- *Post-emergenza*, ovvero discussione di eventi da poco trascorsi che hanno creato situazioni di emergenza, spesso in relazione alla gestione delle attività di ripristino della normalità.

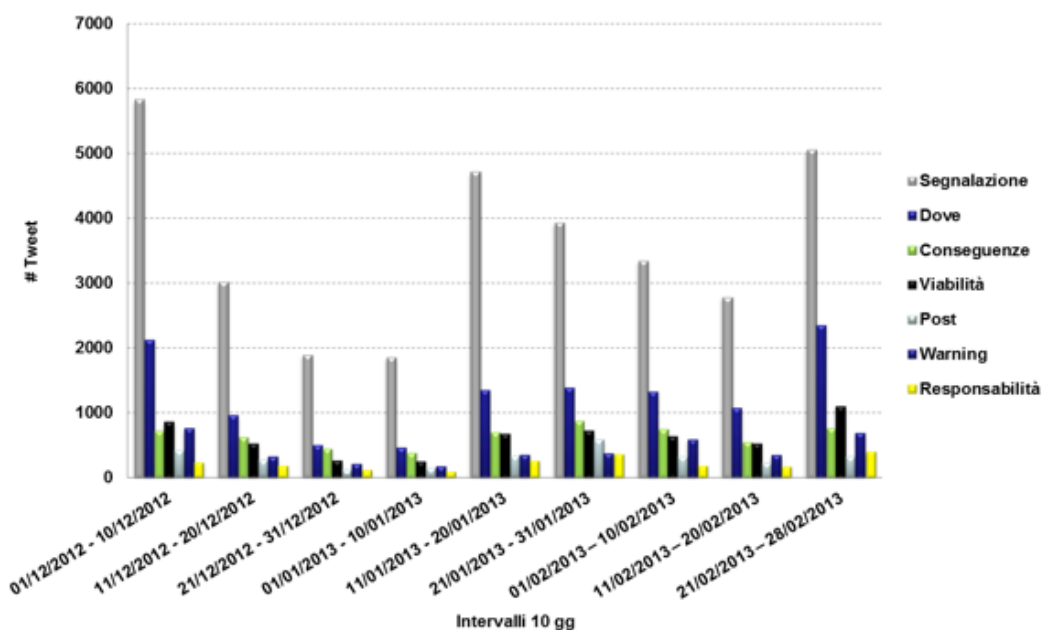


Figura 3
Gli argomenti delle conversazioni online (dominio alluvioni).

La Figura 3 riporta la distribuzione del parlato sugli argomenti di conversazioni sopra descritti. Si nota una prevalenza di attenzione per la categoria *segnalazione* che, a sua volta, indica un ruolo dominante del parlato durante l'emergenza, piuttosto che nelle fasi che precedono e che seguono l'emergenza. Il fatto che le fasi di pre- e post-emergenza siano le meno discusse è parzialmente in relazione a una scarsa presenza delle autorità su Twitter, in

particolare in fase di pre-emergenza. Si nota inoltre un notevole volume di parlato appartenente alle categoria *dove*, ovvero post utili a definire l'area geografica interessata dall'emergenza e *viabilità*, ovvero post che forniscono indicazioni sulle strade interessate dall'emergenza.

Queste osservazioni indicano che:

- La raccolta e analisi delle informazioni da social media, Twitter in particolare, deve essere operata in tempo reale, data la prevalenza di parlato utile in fase di gestione dell'emergenza (si tratta di alcune ore).
- Le informazioni raccolte possono fornire indicazioni utili ai fini delle localizzazione dell'emergenza e, in misura minore, della identificazione di problemi di viabilità.

3. Lo strumento semantico

Tuttavia, non tutta l'informazione si è rivelata utile ai fini della gestione delle emergenze. Come in molti altri domini, oltre metà dei post raccolti sintatticamente tramite parole chiave di crawling si rivela non pertinente rispetto al dominio di analisi. La Figura 3 mostra come il dominio delle alluvioni non faccia eccezione, indicando come circa il 50% dei post raccolti non siano pertinenti al dominio di riferimento pur includendo una delle parole chiave di crawling riportate nella Figura 2. Dati i volumi di parlato, non è possibile pensare a una eliminazione manuale delle informazioni non pertinenti. Nell'ambito dell'architettura progetto Torcia (vedi Figura 1) è il motore semantico [2] ad occuparsi del riconoscimento e dell'eliminazione dell'informazione non rilevante.

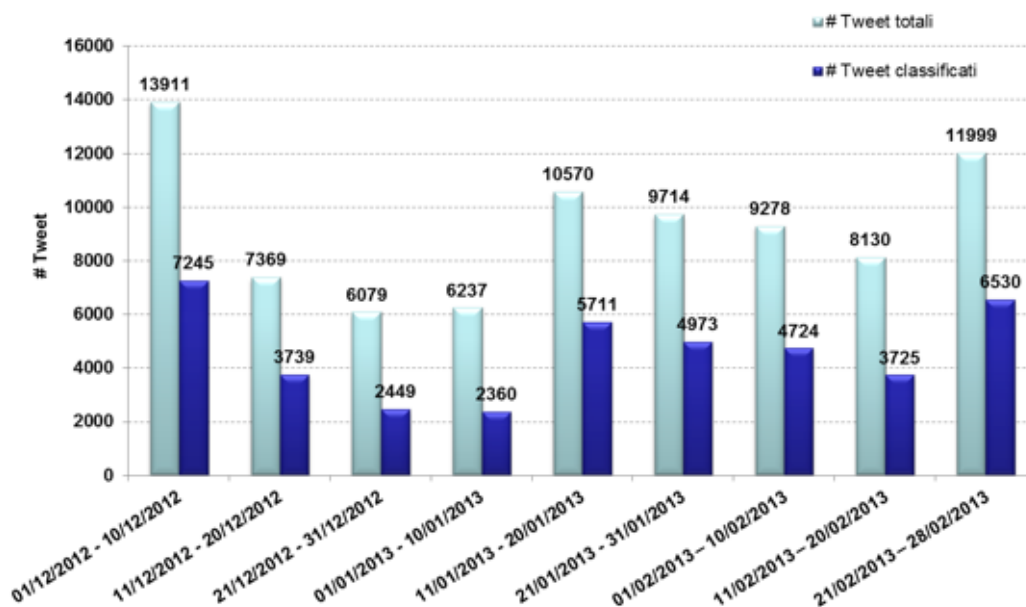


Figura 4
Volumi di parlato raccolti sintatticamente vs. volumi di parlato pertinenti al dominio delle alluvioni (esempio).

Oltre a ripulire l'informazione, il motore semantico classifica l'informazione nelle categorie di parlato precedentemente discusse. Tale classificazione è operata da un modulo software basato su una rete semantica che associa a ciascuna categoria dei metadati, pesandoli opportunamente rispetto al dominio di analisi. La selezione dei metadati e la regolazione dei pesi nella rete semantica sono stati oggetto di attenta analisi e continuo affinamento nel corso del progetto.

Il motore semantico si occupa poi di riconoscere se nell'informazione filtrata c'è l'indicazione di qualche situazione di potenziale emergenza o di emergenza in atto (nel primo caso viene creata un'*alerta*, nel secondo una *segnalazione*). L'individuazione di una situazione di potenziale emergenza, come pure quella di un'emergenza in atto avviene tramite l'individuazione di picchi nei volumi di parlato rispetto ai volumi medi. Tuttavia, si è osservato che non tutti i picchi corrispondono a una reale situazione di emergenza, potenziale o in atto e che non è quindi sufficiente fissare soglie sui volumi di parlato per creare corrispondenti allerte o segnalazioni. Ad esempio, una situazione di maltempo generalizzato su una porzione ampia del territorio nazionale genera un aumento anche molto consistente dei volumi di parlato, ma molto spesso non corrisponde a un'emergenza. Il picco nei volumi è confrontabile o addirittura superiore a quello che si è avuto durante l'alluvione a Catania del 2013 o quella in Sardegna del 2014. Ciò che permette l'individuazione di una situazione di potenziale emergenza o di emergenza in atto è la combinazione di soglie di parlato con una percentuale minima (una terza soglia) di volumi che si riferiscono a una stessa zona o località geografica (categorie *dove* e *viabilità*). Ad esempio, nel caso dell'alluvione a Catania, ciò che il motore semantico scopre analizzando i dati è un picco sopra soglia di parlato che fa contemporaneamente riferimento al dominio *alluvioni* e alla località geografica *Catania*. Il riconoscimento delle località geografiche avviene tramite interfacciamento con il servizio cartografico realizzato all'interno del progetto da Vidiemme (vedi Figura 1).

Categorie	Metadati che le compongono
Dove	Città, Centro, Via, Metro, Nord, Spagna, Venezia, Posto, Località, Zona...
Segnalazione	Allagamenti, Pioggia, Esondato, Fiume, Acqua, Tempesta, Nubifragio...
Conseguenze	Morti, Salvo, Frana, Evacuata, Inghiottiti, Situazione, Danni, Crollata...
Responsabilità	Polizia, Magistrato, Pattuglia, Autopompa, Ministri, Comune, Tilt, Capo...
Viabilità	Libera, Percorribile, Ripulita, Sgombra, Chiusa, Traffico, Percorsi, Frana...
Warning	Fogne, Intasato, Argini, Pericoloso, Previsti, Rischio, Allerta, Flusso...
Post	Risarcimento, Lamentela, Soldi, Assegni, Fondi, Denunce, Interventi...

Figura 5
Categorie di parlato e esempi di metadati.

Il cruscotto direzionale

Allerte e segnalazioni vengono passate al cruscotto di monitoraggio (realizzato da Beta80, partner di progetto). Nel corso del progetto, abbiamo avuto modo di verificare l'architettura Torcia su molte situazioni di emergenza, fra cui:

- alluvione Catania Febbraio 2013,
- alluvione Vicenza Maggio 2013,
- (primo) nubifragio Roma, 07/07 - 08/07, 2013,
- (secondo) nubifragio Roma, 21/07 - 22/07, 2013,
- (terzo) nubifragio Roma, 27/08 - 28/08, 2013,
- allagamenti Toscana, 21/10 - 22/10, 2013,
- alluvione Sardegna Novembre 2013,
- alluvione Marche Aprile 2014,
- alluvione Senigallia Maggio 2014,
- esondazione del Seveso Settembre 2014.

Ciò di cui ci siamo resi conto è che il parlato su social media è di per sé poco fruibile se non viene organizzato, ovvero ripulito, categorizzato e geolocalizzato. Questo sforzo di analisi permette però di ottenere informazioni interessanti a complemento di quelle provenienti da altre fonti già tradizionalmente utilizzate dagli operatori istituzionali. La Figura 6 mostra, nel caso dell'alluvione a Catania, come le allerte sollevate dal motore semantico sono inviate al cruscotto decisionale con un anticipo temporale significativo rispetto al momento dell'allerta ufficiale. Questo comportamento è consistente su tutte le emergenze analizzate.

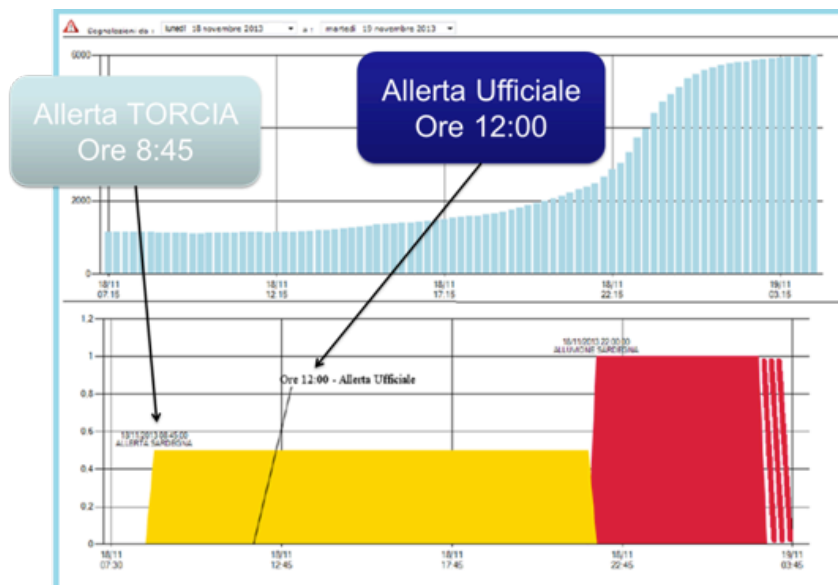


Figura 6
I tempi delle allerte TORCIA.

Tuttavia, l'approccio con il quale è stato progettato il cruscotto di monitoraggio è fortemente orientato all'operatore, al quale le informazioni, incluse allerte e segnalazioni, sono soltanto proposte e devono essere validate manualmente. Tale validazione è stata ritenuta necessaria sia perché il motore semantico, nonostante il continuo affinamento, resta comunque soggetto a errore, sia perché non tutta l'informazione sociale è affidabile e deve essere analizzata e validata da un operatore che ha accesso a tutte le possibili fonti informative, non solo ai social media.

La Figura 7 mostra la schermata del cruscotto con gruppi di tweet geolocalizzati nel caso dell'alluvione in Sardegna del novembre 2013. Si nota come l'informazione resa disponibile dal cruscotto sia molto più fruibile per un operatore istituzionale rispetto all'accesso diretto ai social media.

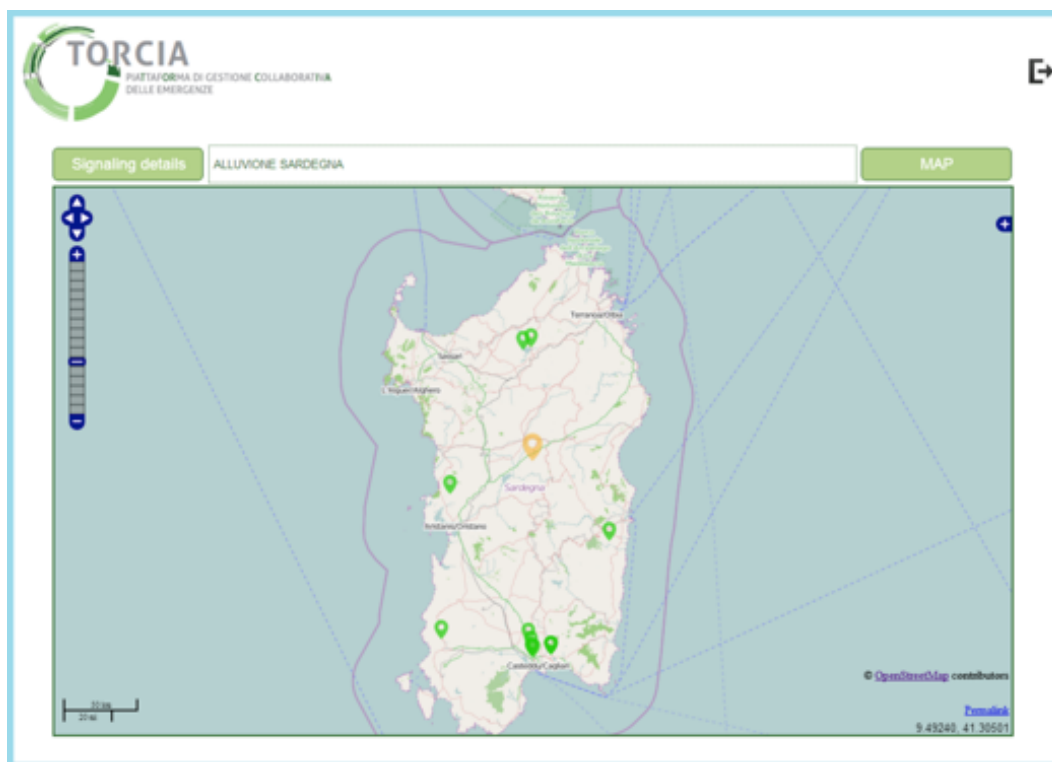


Figura 7

Geolocalizzazione delle informazioni da social media: il caso dell'alluvione in Sardegna.

4. Formazione informatica e produttività degli utenti

L'informazione organizzata e validata dagli operatori è resa disponibile ai cittadini tramite un'applicazione mobile (realizzata da Beta80). La applicazione mobile rappresenta il punto di incontro fra cittadini e operatori istituzionali, costituendo uno strumento di coordinamento e, per le istituzioni, di crowdsourcing. I cittadini possono accedere a informazione validata e,

direttamente dalla app, fornire ulteriore informazione agli operatori, inviandola direttamente al cruscotto di monitoraggio e contemporaneamente anche a Twitter. Gli operatori possono raccogliere informazioni dai cittadini, in tempo reale, e a loro volta inviare comunicazioni importanti per la gestione delle emergenze. Inoltre, tramite il meccanismo di registrazione, i cittadini utenti della mobile app rappresentano un insieme di utenti privilegiati, dei quali si può conoscere l'identità e tracciare la storia in caso di emergenze ricorrenti. E' poi possibile valutare la significatività dei contributi degli utenti della app, con i noti meccanismi di scoring degli utenti, in modo da poter operare valutazioni anche automatiche di affidabilità delle informazioni sulla base degli score.

La Figura 8 mostra l'accesso alle informazioni geolocalizzate tramite la mobile app. E' evidente come gli stessi operatori di protezione civile possano essere utenti della mobile app [1] e come, di conseguenza, la app rappresenti uno strumento di coordinamento fra cittadini e operatori istituzionali. E' possibile immaginare scenari di gestione delle emergenze in cui la applicazione mobile venga utilizzata anche per selezionare cittadini e assegnare loro un ruolo nella gestione dell'emergenza, utilizzando le informazioni dei profili utenti e la loro storia di utilizzo della app.

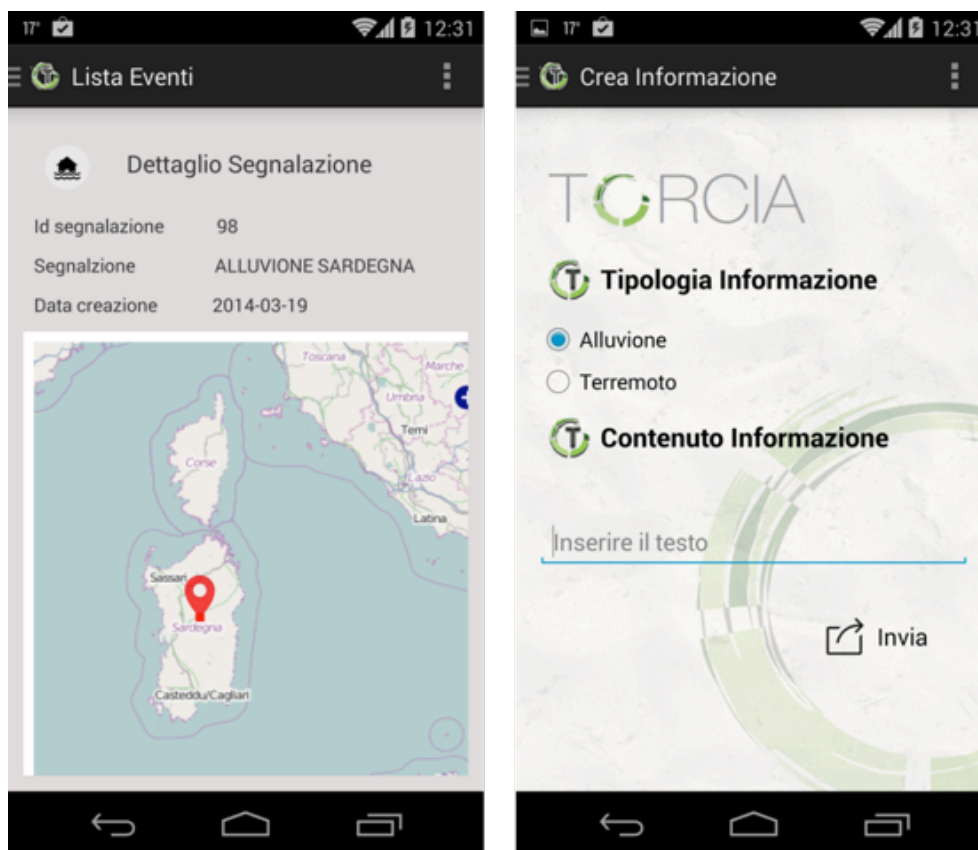


Figura 8
Geolocalizzazione delle informazioni da social media con l'applicazione mobile.

5. Conclusioni

Riassumendo, il workflow di TORCIA parte con la raccolta in tempo reale di informazioni dai principali social media. Tale informazione viene analizzata da un motore semantico che è in grado di individuare e localizzare geograficamente situazioni di potenziale emergenza in base al parlato online. A seguito di tali situazioni, viene creata un'allerta che è poi inviata ad un cruscotto di monitoraggio progettato per gli operatori istituzionali, quali la Protezione Civile. Se l'allerta viene convalidata dagli operatori, si crea una segnalazione di emergenza e si attivano tutte le procedure di gestione dell'emergenza. In particolare, all'interno della piattaforma TORCIA, è possibile comunicare con i cittadini in modalità multi-canale, includendo nello sforzo di comunicazione anche gli stessi social media. I cittadini possono utilizzare la mobile app per ottenere informazioni utili sulla situazione di emergenza. Ad esempio, possono accedere a tutta l'informazione che il motore semantico preleva dai social media visualizzandola su una mappa geografica e possono richiedere alla app il calcolo delle vie di fuga in base alla posizione corrente e alle indicazioni operative degli operatori istituzionali (questo modulo è stato sviluppato da ACT Solutions). Tramite la app possono a loro volta fornire informazioni utili per la gestione dell'emergenza, favorendo la realizzazione di un processo collaborativo che integra i contributi di cittadini e istituzioni.

Ringraziamenti

Gli autori desiderano ringraziare le aziende partecipanti al progetto TORCIA e, in particolare, Claudio Mangone e Giorgio Parladori per il loro contributo scientifico e di coordinamento.

Bibliografia

- [1] Capelli, P. "Una metodologia per la classificazione dei modelli competitivi delle mobile application", Tesi di Laurea in Ingegneria Informatica, Politecnico di Milano, Dicembre 2013.
- [2] Carcaci, C. "Classificazione sintattica con metodi insiemistici", Tesi di Laurea in Ingegneria Informatica, Politecnico di Milano, Dicembre 2012.
- [3] Cesana, G., D'Etto, A. "Analisi, Metodologia e Sviluppo di un Sistema Integrato per la Gestione delle Emergenze in Ambito Social", Tesi di Laurea in Ingegneria Informatica, Politecnico di Milano, Luglio 2013.
- [4] Rossitto R., "I numeri di Twitter in Italia", Wired, 26 Settembre 2012, <http://daily.wired.it/news/internet/2012/09/26/numeri-twitter-italia.html>
- [5] Lachlan, K.A., Spence, P.R., Lin, X., Del Greco, M. "Screaming into the Wind: Examining the Volume and Content of Tweets Associated with Hurricane Sandy," *Communication Studies*, 65(5), 500-518, 2014