A Screening of New Media Tool for Crisis Management

Joel Brynielsson, Fredrik Johansson, Carl Jonsson, Sinna Lindquist, Maribel Narganes Quijano, Anders Westling

FOI

Outline

- Introduction
  - Crises and social media
  - Alert4All and the Screening of New Media (SNM) tool
- Methodology
  - User-centered tool design
  - Creating appropriate machine learning classifiers
  - Experiments and results
- Tool overview

Crises and social media

- The use of social media during crises is becoming increasingly common.
  - Fukushima (see [Doan et al., 2011] and [Thomson et al., 2012])
  - The terror attacks in Norway (see [Perng et al., 2012])
  - The San Bruno fire (see [Nagy and Stamberger, 2012])

"Country residents outside of Fargo are surrounded by flood waters."

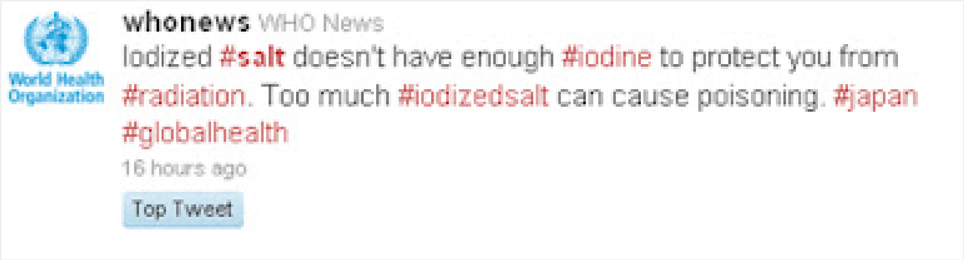
"@NilsPetter We are sitting by the lake. A man dressed in police uniform is shooting. Help us regarding when the police will arrive."

- Most current research is concerned with retrospective analysis of the use of social media in crises.
  - More and more attempts are made to make use of social media for improved situation awareness **during** the crisis.

Crises and social media

- Parts of the User Generated Content (UGC) in social media can be used for an increased tactical situational awareness:
  - More data than can be processed manually
  - A low signal-to-noise ratio (much uninteresting and unreliable information)

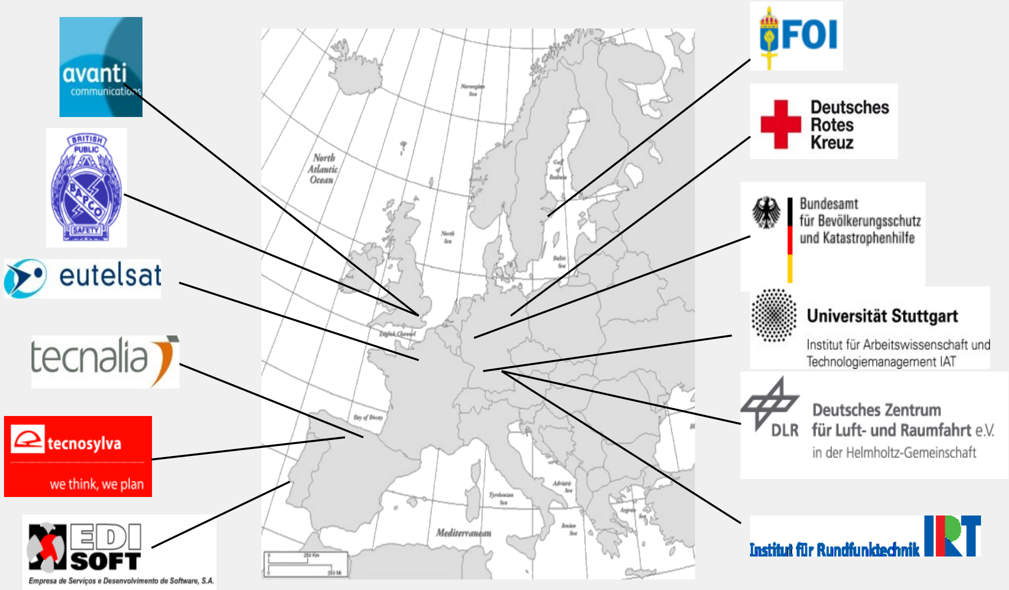
- Need for automatic or semi-automatic techniques for supporting crisis management!



FOI

Alert4All

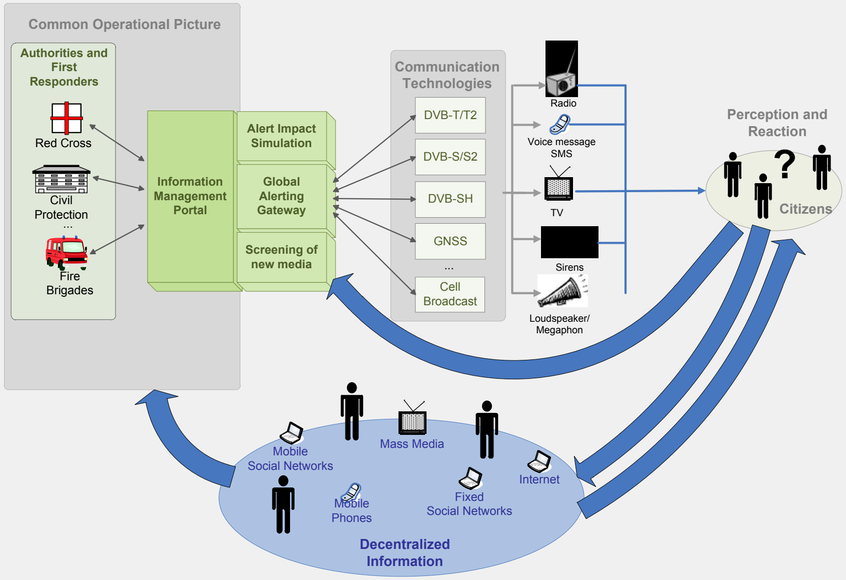
- Alert4All is a EU-funded FP7 security project : 2011 – 2013
- Focus: to improve the effectiveness of alert and communications to the population in contemporary crises
- Consisting of a team of 12 European partners



FOI

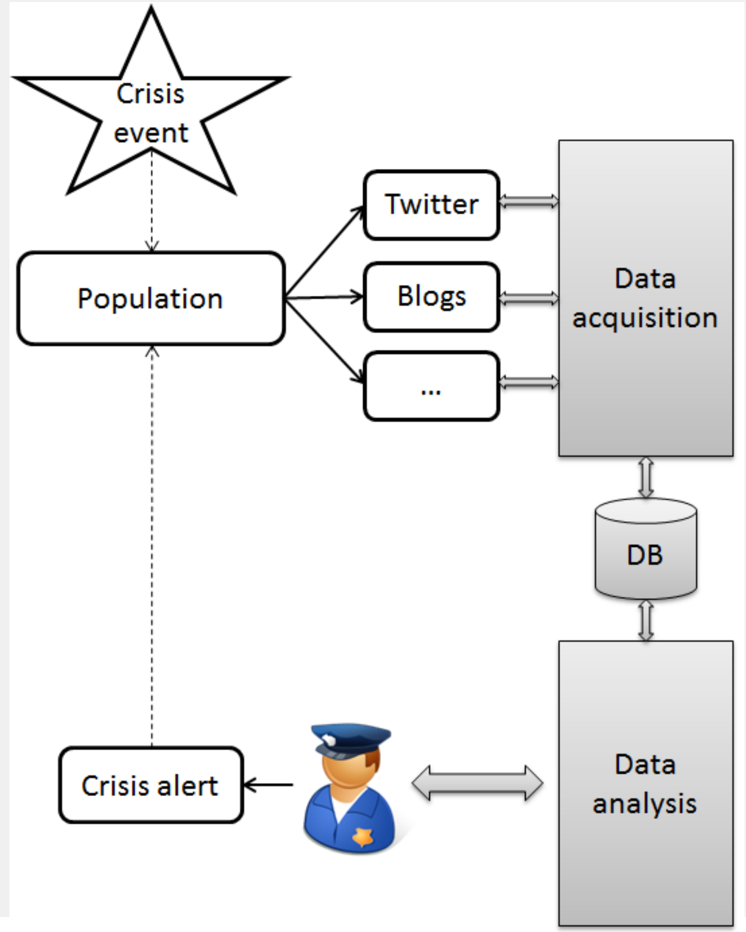
The Alert4All Concept

- Development of a portal for information share and management to allow cooperation between authorities
- Multi-channel approach to maximise the number of reached citizens
  - Employing harmonised best practices to compose alerts
- Awareness of alert impact
  - ‘What-if’ simulation tool considering human behaviour
  - **Screening of social media in real time**



The Screening of New Media (SNM) Tool

- We want to find out how citizens react to the crisis and communicated alert messages
- Assumption: parts of the population make use of social media during the crisis
- Collection of UGC from various sources
  - Twitter
  - Blogs
  - YouTube



User-centered tool design

- Workshops with end users and stakeholders:
  - Pre-study
  - Design workshop – video prototyping
  - Validation exercise

Design issues

- What kinds of emotional states are relevant to consider operationally, i.e. what public emotions do crisis management personnel (the users) consider to be important to know about in order to make an informed decision in a crisis situation?
- How would the users prefer to have the result of the social media screening presented?

FOI

Design principles

- User centeredness
- To go from the specific to the generic
- Make use of end users’ and other stakeholders’ domain expertise and work skills, all experts in their respective field
- Usable methods for
  - framing the field
  - including users on equal terms
  - generating design ideas as well as generating potential solutions for those ideas
- Future use of future technology (much is uncertain)

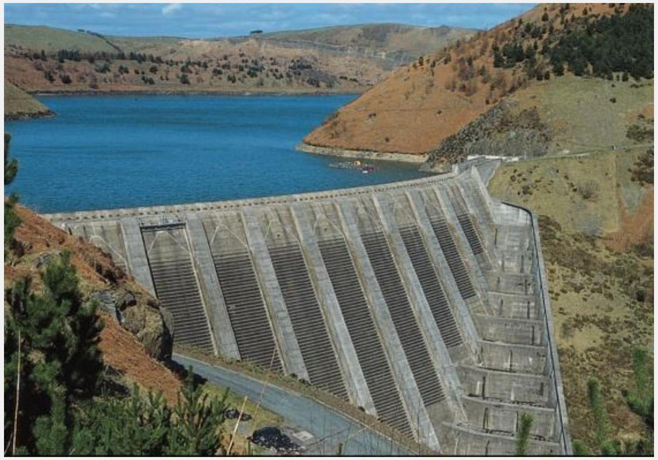
End users and stakeholders

- Crisis management personnel
- Researchers, developers, programmers representing different
  - organizations,
  - European countries,
  - positions, languages, and cultures,
  - communicative and technical skills and possibilities.

Video prototyping

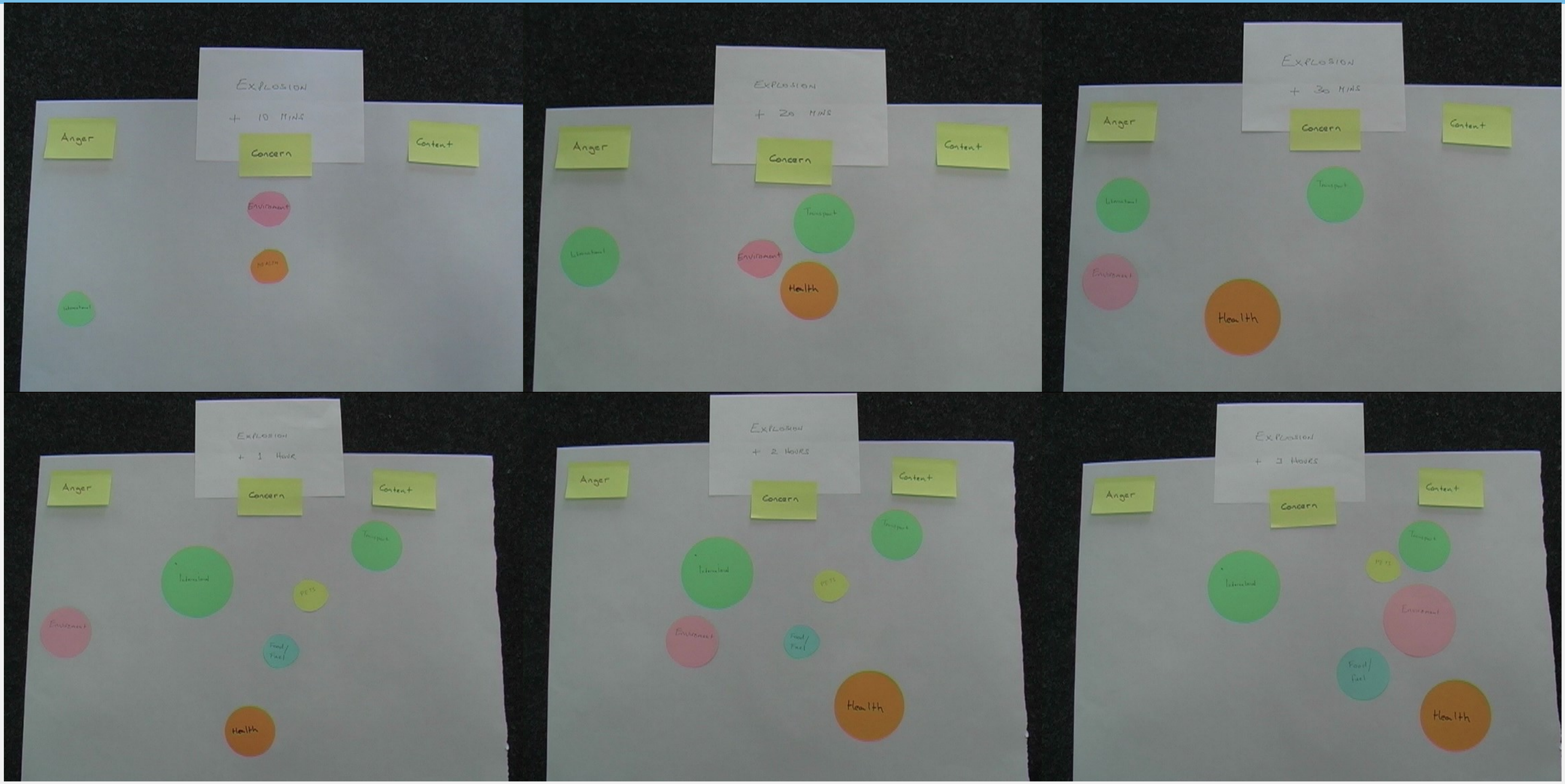
- Video prototyping
  - Methodology to illustrate and communicate design ideas
  - Based on real life experiences and end user knowledge
  - Explicit, not general
  - Focused, not everything
  - Show and tell
  - Shared understanding of design problems and design ideas
  - “Quick and dirty,” i.e., no editing (and no Oscar award chances!)

Video prototyping workshop layout



- Presentation of the SNM tool idea
- Description of scenarios
- Introduction to video prototyping
- Brainstorming post-it session, clustering, naming clusters
- Idea-generation/sketching:
  How would you want to get the information presented?
- Shooting of the video-prototype
- All watching all video-prototypes, synthesizing and drawing conclusions for further design

Video prototype result



Video prototyping results

- Most important: in a crisis situation one wants to bring a situation into a, in some sense, better situation
- “Acting on facts, not emotions,” but situation awareness includes emotions
- Quantification of emotions and attributes
- Negative emotions, such as “fear,” and “anger,” are more important to consider and distinguish between than positive emotions such as “happiness” and “relief”

Developing classifiers for affect analysis

- Issues related to development of machine learning-based affect analysis of crisis-related posts:
  - Should a post be classified as *positive*, *fear*, *anger*, or *other*?
    - Those classes is a result from workshops held with crisis responders

- Most previous research has focused on sentiment analysis (positive / negative)
  - Affect analysis is a harder (multi-classification) problem
  - Novel contribution: affect analysis of posts for the crisis management domain

Collecting relevant tweets

- Used Python package **tweetstream** to retrieve tweets related to the Sandy hurricane
  - Random sample of the total volume of tweets
  - More than 6 million tweets received using:
    - **Sandy**, **hurricane**, and **#sandy**
- After removal of re-tweets, duplicated tweets and non-English tweets approximately 2.3 million tweets remained
  - Not feasible to annotate the whole dataset manually…
  - Large proportion belonging to class *Other* (non-emotional content)

Creating a training set

- Step 1: Create a set of tweets likely to contain relevant emotions
  - Seed-words based on manual inspection
    - Angry, furious, hate, afraid, scared, glad, happy, :), =), etc.
  - Automatically extended using synonyms from WordNet
  - Used for sampling 1000 tweets for each tag (”positive,” ”fear,” …)
    - Biased selection process!
    - Risk of ending up with classifiers that only learns the words used for sampling
    - Only solution we could identify for creating a balanced training set without requiring a huge number of annotators
    - As we will see later, the classifiers have learned to generalized beyond the used simple rules

Creating a training set

- Step 2: Use human annotators for labeling
  - Team of 12 independent annotators connected to the Alert4All project
  - Each annotator received a file with 1000 tweets
  - Task: Annotate each tweet with the correct class label
    - Correct class label decided through majority agreement

Category	Majority agreement	Full agreement
"positive"	92.7%	47.8%
"anger"	92.6%	39.2%
"fear"	95.2%	44.4%
"other"	99.7%	82.3%

TABLE I
INTER-ANNOTATOR AGREEMENT FOR THE VARIOUS CATEGORIES

Emotion class	Number of tweets
Positive	622
Anger	461
Fear	470
Other	2249

TABLE II
NUMBER OF TWEETS PER CLASS (BASED ON MAJORITY AGREEMENT)

- Step 3: Creation of balanced training set
  - 461 training samples for each class

Creating a separate test set

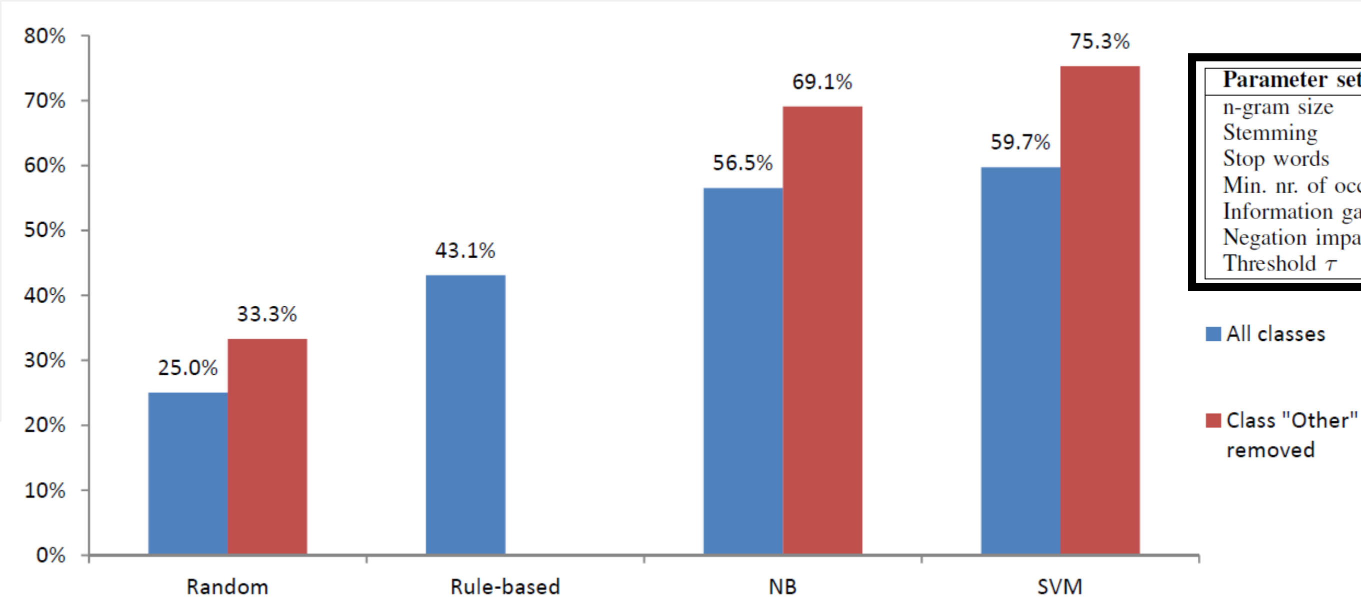
- Use of separate test set instead of using cross-validation
  - Motivation: avoiding eventual biases
- A human annotator classified sampled tweets until we got a sufficiently large test set
  - 54 tweets in each class (after balancing the classes)
  - Training set, approx. 90 %
  - Test set, approx. 10 %

Experiments

- Classifiers
  - Naïve Bayes (NB) and Support Vector Machines (SVM)
- Features
  - Bag-of-words representation
  - Many parameter settings have been tested
    - N-gram size: 1 (unigram) / 2 (unigram + bigram)
    - Stemming: yes / no
    - Stop words: yes / no
    - Min. nr of occurrences: 2 / 3 / 4
    - Information gain (in %): 25 / 50 / 75 / 100
    - Negation impact (nr of words): 0 / 1 / 2
    - Threshold : 0.5 / 0.6 / 0.7

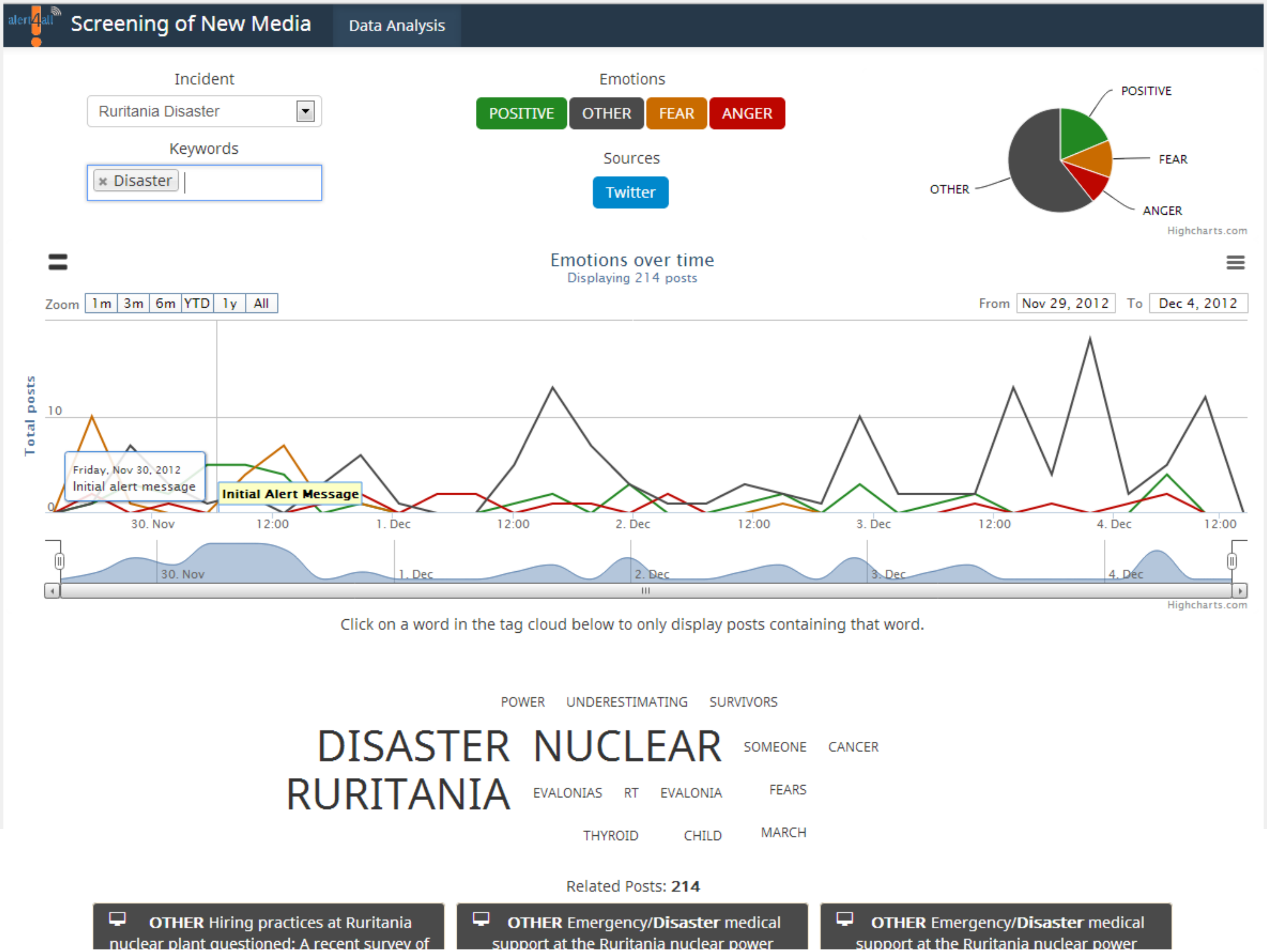
Experimental results

- Results are much better than the baselines
  - The classifiers have learned something useful (not only the keywords)
  - Still, probably not good enough for classification on the level of individual tweets in a real-world crisis management system
    - Good enough to be used on an aggregate level?
  - Hard task as suggested by the low full agreement among annotators



Parameter settings	SVM	NB
n-gram size	2 (unigram + bigram)	1 (unigram)
Stemming	yes	yes
Stop words	yes	yes
Min. nr. of occurrences	4	4
Information gain	75%	75%
Negation impact	2	2
Threshold τ	0.7	0.6

The Screening of New Media (SNM) Tool



Summary: design issues

- Understanding overall emotions and attitudes among the population is likely to be beneficial for crisis communication
- Important aspects in order to enhance situational awareness:
    - Variation of public emotions and attitudes over time, specifically in relation to sent alert messages
    - Quantification of emotional states
    - Distinguishing between different negative emotions
    - Adaptive GUI that allows for several kinds of representations and visualizations

FOI

Summary: algorithm development

- A methodology for classifying large amounts of crisis-related posts has been devised
- A first attempt to find emotions in crisis-related tweets give promising results
  - Accuracy ~60 % for the SVM classifier
- Future work to be evaluated on other datasets
  - How well does the classifiers generalize to other kinds of crises?

Questions?

Thank you for your attention!