



Project started in 2013



AIDR: Artificial Intelligence for Digital Response

Qatar Computing Research Institute

Muhammad Imran (Project & Science lead)

Ferda Ofli (Computer Vision Science lead)

Firoj Alam (Postdoc)

Partners & Customers:



New York (Suffolk)
Emergency Management Dept.



Time-Critical Events and Information Gaps

Disaster event (earthquake, flood)



Urgent needs of affected people



- Food, water
- Shelter
- Medical assistance
- Donations
- Service and utilities

Information gathering,
especially in real-time, is
the most challenging part



Information gathering



Relief operations



Humanitarian organizations and local administration



OCHA



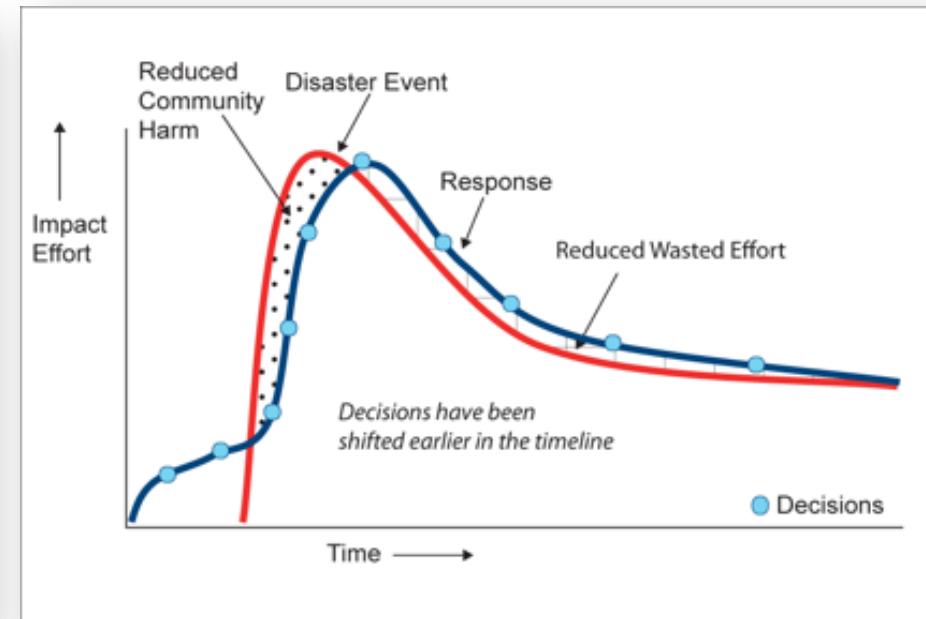
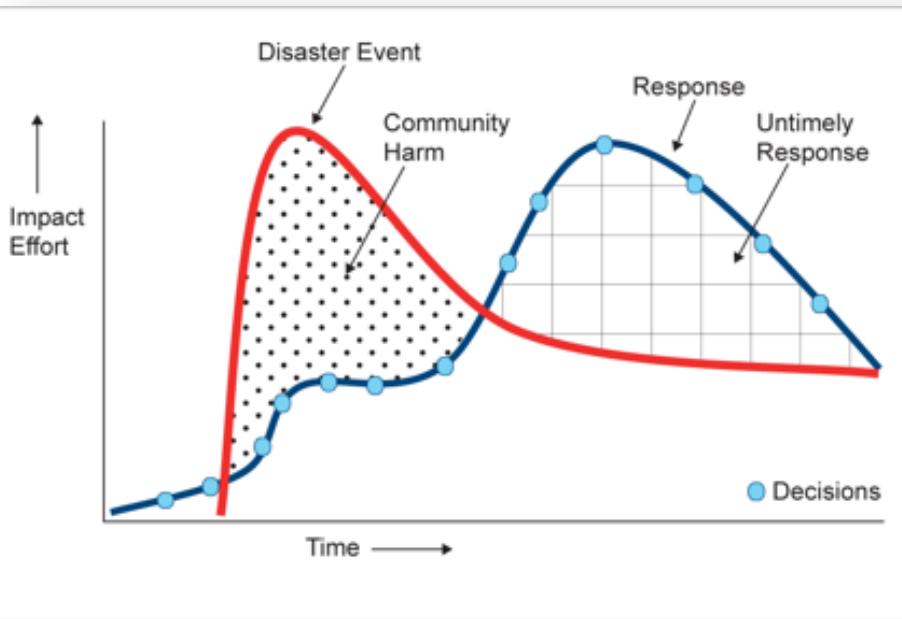
unicef



World Health Organization



Decision-making and Response

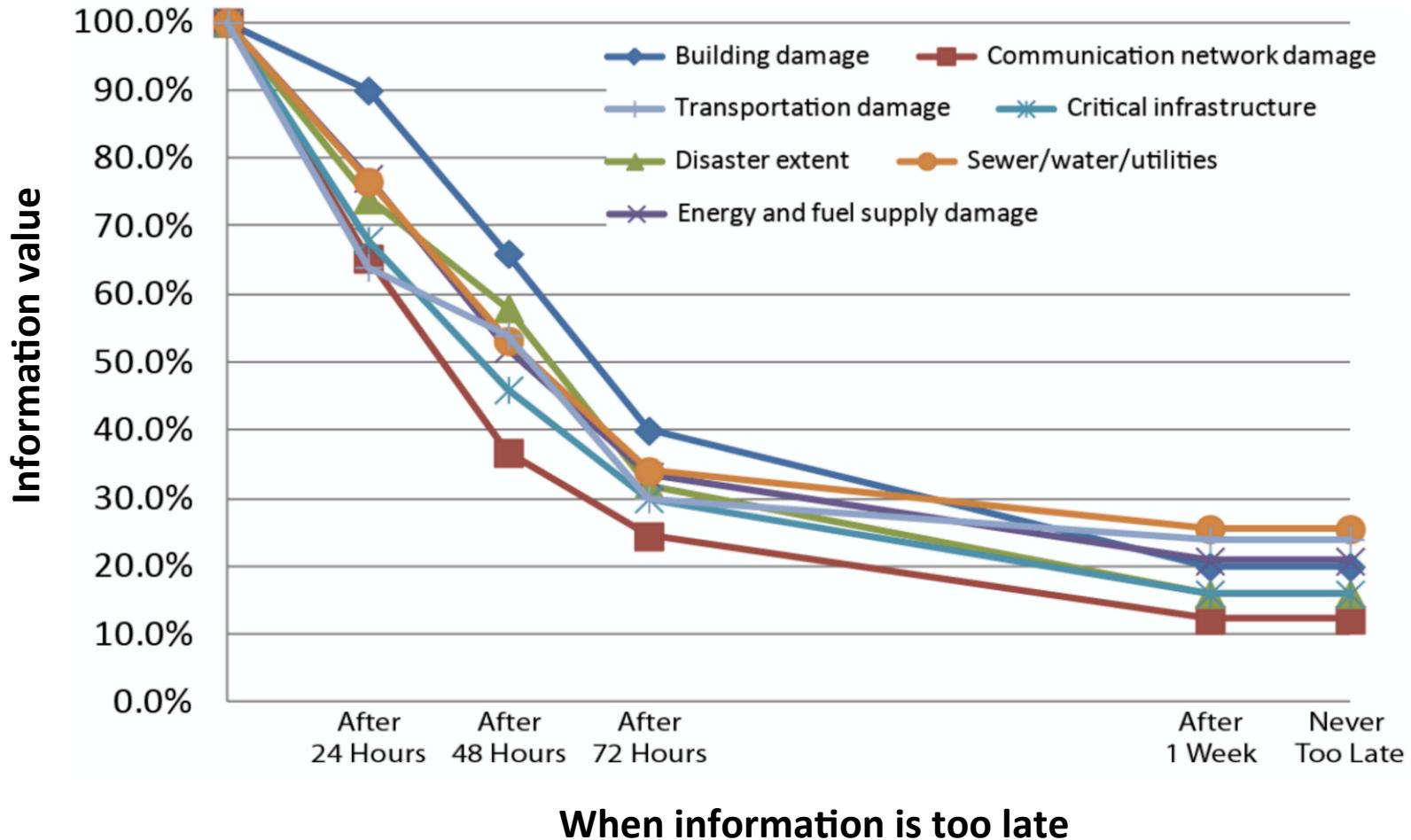


- **Delayed decision-making**
- **Delayed crisis response**
- **High community harm**

Target

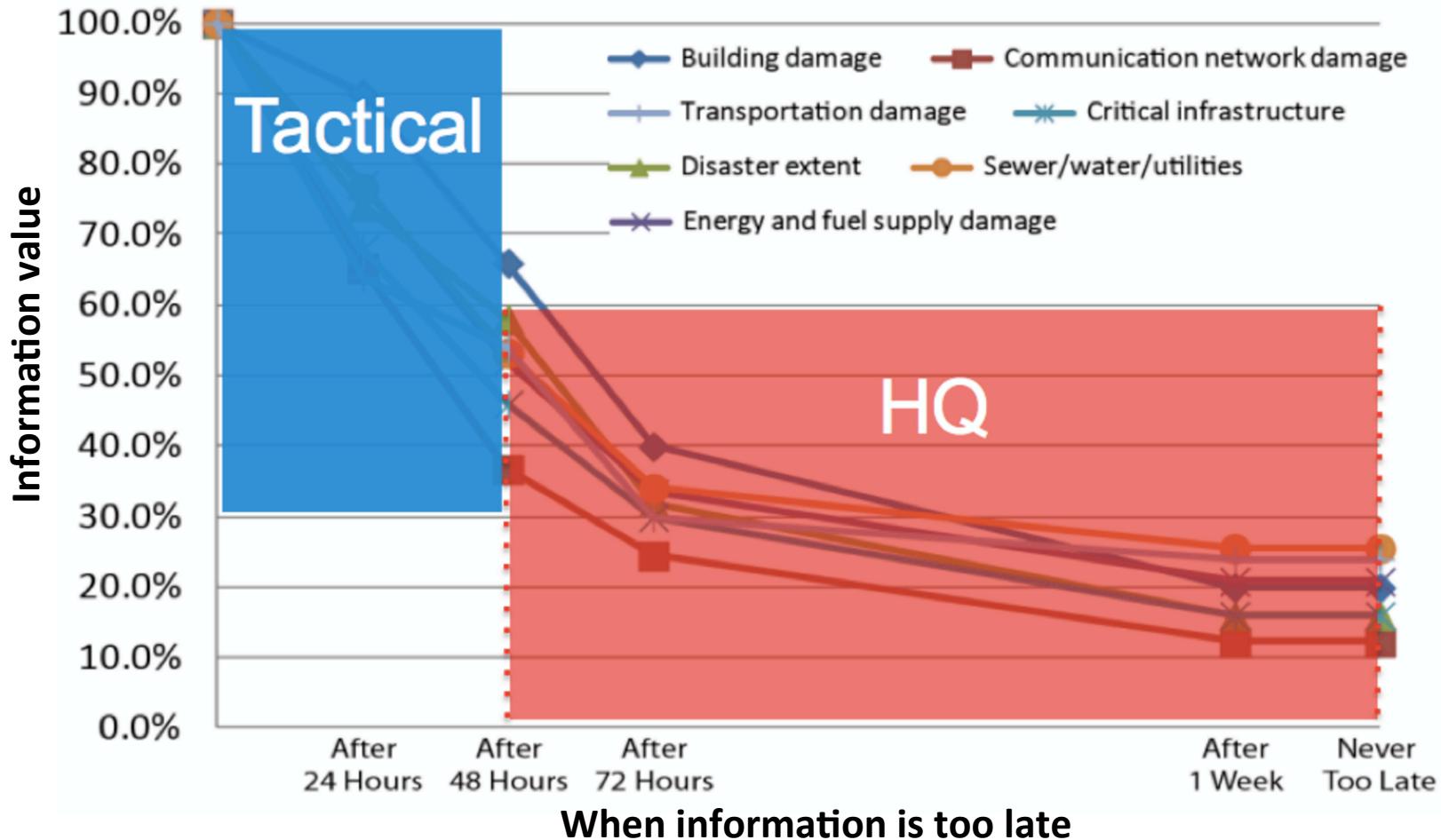
- **Early decision-making**
- **Rapid crisis response**
- **Low community harm**

The Value of Timely Information



Based on FEMA large-scale survey among emergency management professionals across the US.

The Value of Timely Information



Based on FEMA large-scale survey among emergency management professionals across the US.

Current Customers and Uses



To gain situational awareness and to extract actionable information



To answer HIV/AIDS-related health questions using SMS



To find suspicious reports, road accidents, blockage, power outages etc.



To find reports of attacks and destruction of cultural heritages sites in Yemen, Syria



To develop effective solutions for disaster management



Consumes and triangulates important information for small response organizations

Typhoon Hagupit (2014)

the guardian
Winner of the Pulitzer prize

UK world sport football opinion culture economy lifestyle fashion environment tech money travel

home > world europe US americas asia australia africa middle east

Typhoon Hagupit

Typhoon Hagupit: UN using crowdsourcing platform to help assess damage

The United Nations is working with crowdsourcing platform, MicroMappers, to assess how much destruction has been caused by the latest typhoon to hit the Philippines



© Fishermen dock their boats at Manila Bay as Typhoon Hagupit (locally known as Ruby) approaches on 7 December 2014, Pasay City, Philippines. Photograph: Mark Cristino / Barcroft Media/Mark Cristino / Barcroft Media

At least 27 people have been killed and a million people evacuated after Typhoon Hagupit (locally known as Ruby) made landfall in the Philippines on Saturday night. Though Hagupit appears to be weakening as it approaches the Philippines capital, Manila, it's thought to have destroyed around 80% of all the homes along some coastal areas.

Like Typhoon Haiyan, last year, technology is again being used to help identify damage and needs assessment on the ground. The [United Nations Office for the Coordination of Humanitarian Affairs \(OCHA\)](#) in Manila has requested support

Advertisement AdChoices

amazon web services AWS

Linux or Windows Server FREE for 1 Year Try it now

replay

Most popular



India v West Indies: Cricket World Cup - live!
| Dan Lucas and Simon Burnton

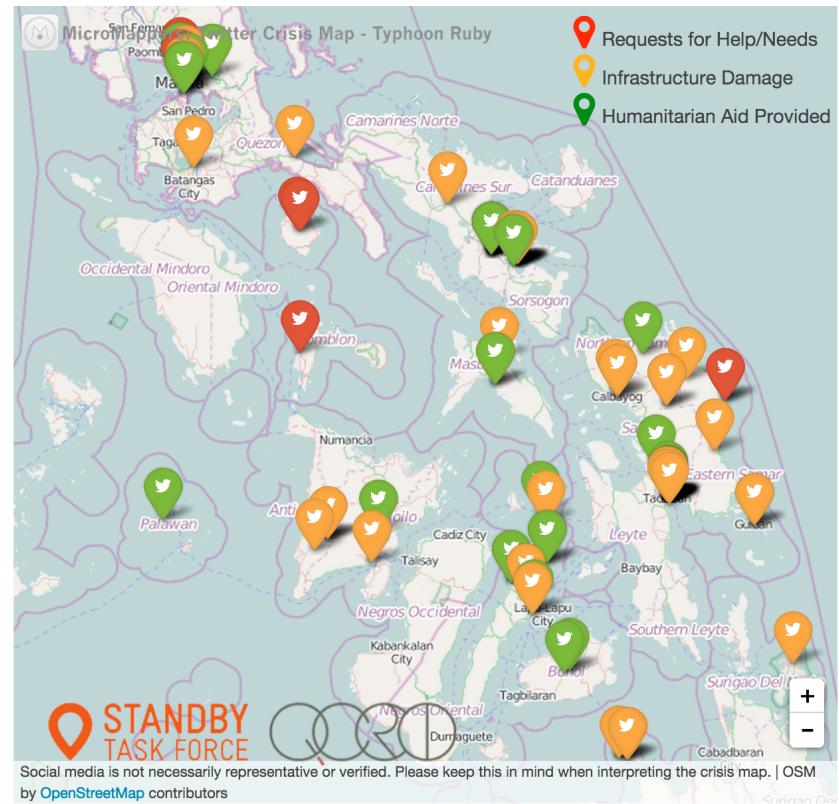


Harrison Ford hospitalised after crashing his plane on golf course



Nasa finds evidence of a vast ancient ocean on Mars

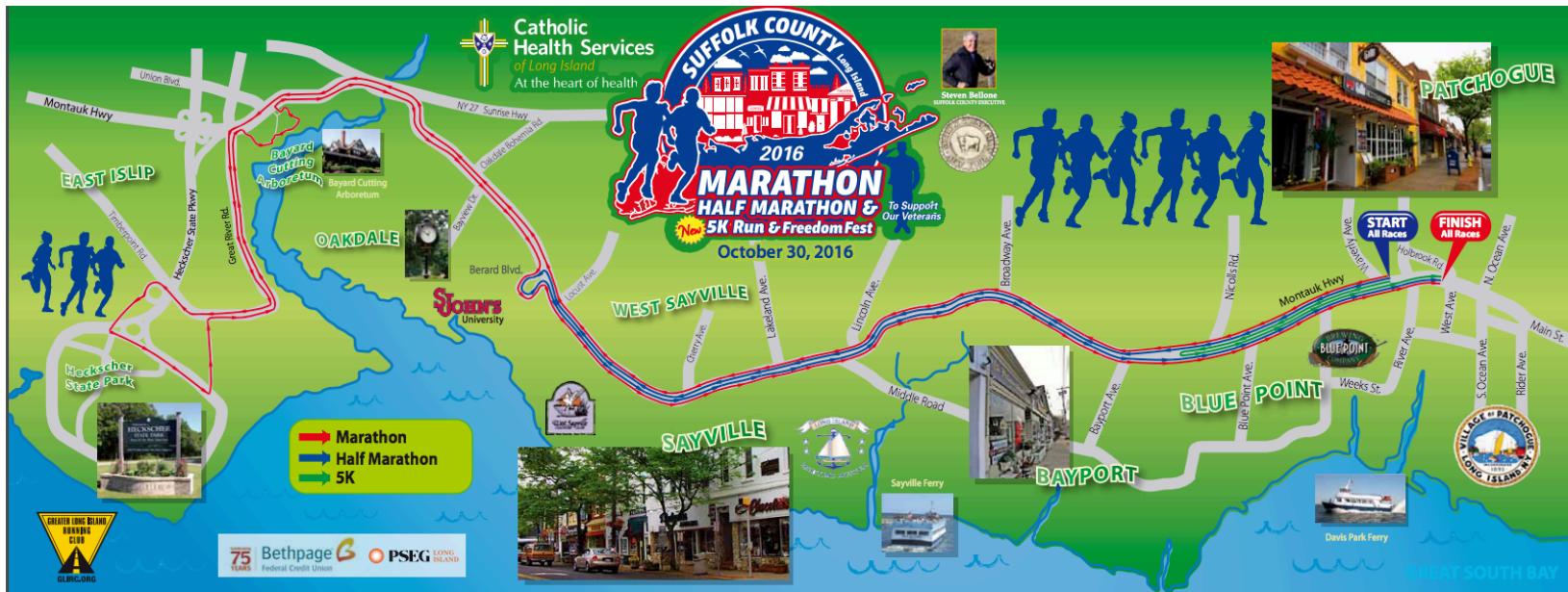
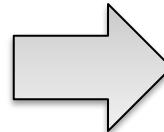
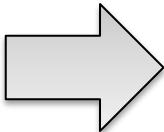
MicroMappers Map showing tweets which mention urgent needs, infrastructure damage and humanitarian aid regarding Typhoon Ruby.



As per the ICRC's guidelines on data protection and privacy, the Tweet Map does not include any personal identifying information (PII). This PII along with the raw tweets are only shared with the United Nations.

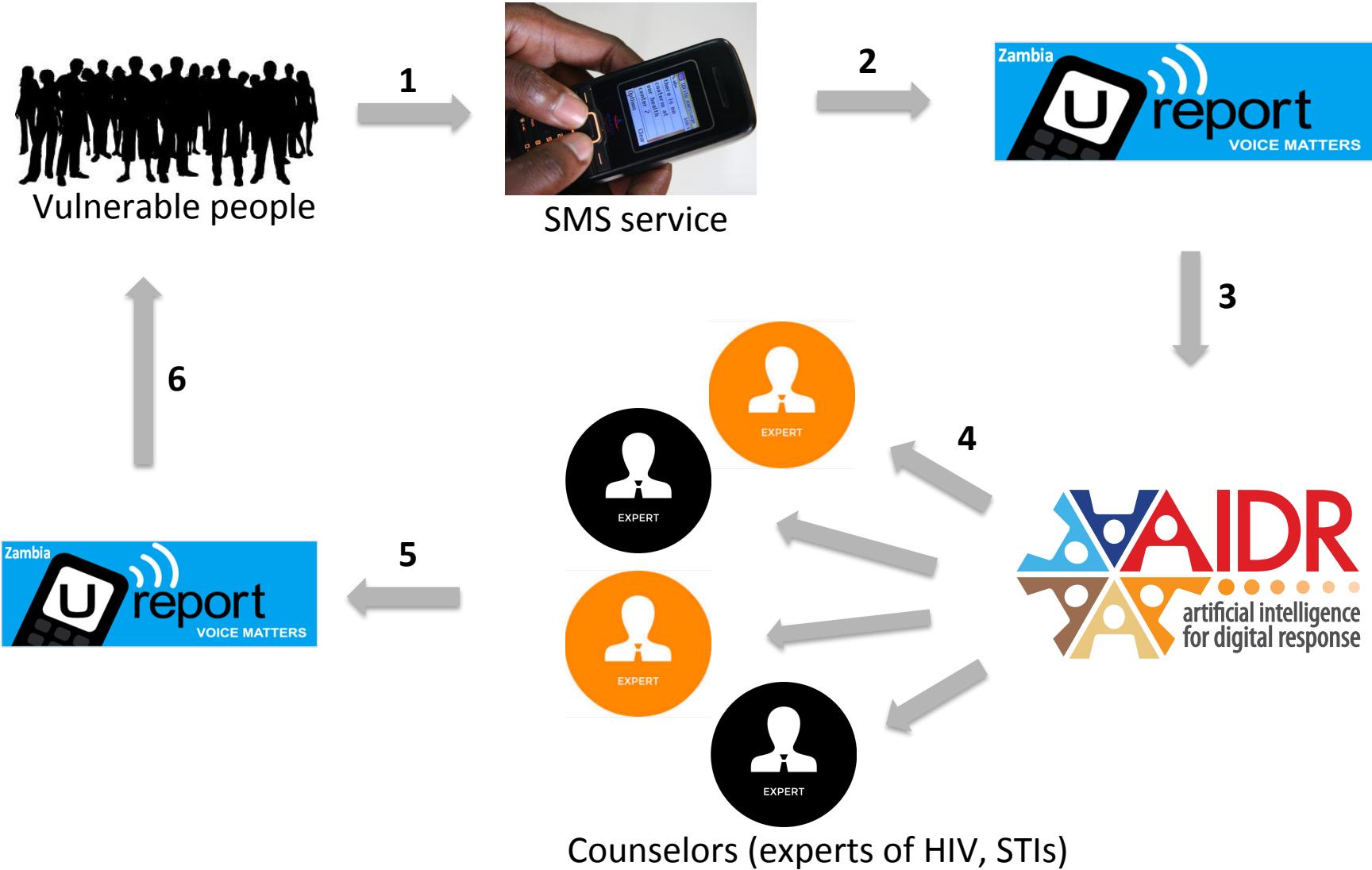
- Killed 27 people
- A million evacuated
- \$114 million of damage

Suffolk Marathon 2015



Task: Suspicious reports/images and accidents detection

AIDR Helps Answer Thousands of Health Queries



New Scientist Featured This Work



DAILY NEWS 4 April 2016

AI helps answer thousands of health queries in Zambia via SMS



SMS has proven an exceptionally useful technology in parts of Africa
GIANLUIGI GUERCIA/AFP/Getty Images

For many people in Zambia with health queries, sending a text message is the best way to get it answered. U-report, a free SMS-based service set up by UNICEF and run by volunteers, receives many thousands of questions a month, many specifically about [HIV and AIDS](#).

Also popular in Uganda, U-report has seen usage triple in the last three years, and about a thousand new users register every day. The volume of messages is growing so fast that the volunteers can't keep up, so UNICEF is testing software that reads and responds to many of the messages automatically.

In Zambia, there are roughly 27,000 new HIV infections a year, according to UNICEF, and 40 per cent of these are in those aged 15 to 24. With people constantly texting U-report for all kinds of HIV information and advice, the automated version [uses machine learning algorithms](#) to sort messages into eight categories: symptoms, HIV testing, treatment, pregnancy, transmission, prevention, definition, and male circumcision.

To train the system, [Patrick Meier](#), then at the Qatar Computing Research Institute in Doha, and colleagues fed in at least 50 messages for each category that had been selected by hand, and asked it to identify patterns that it could then use to do the sorting itself. As well as how to handle typos, the system learned to cope with textspeak such as "HOW 2 AVOID SPREADING HIV/AIDS 2 OTHERS?" and "I feelin bad becoz im th only one wh hs hiv wht shld i do?"

UNICEF Testimonial

UNICEF feedback: “Our overall experience with AIDR-SMS system was satisfactory. Our test found that the AIDR-SMS system offered high accuracy results on SMS that contained primarily a single topic or belonged to single category. This is to be expected and was explained to us by the AIDR-SMS team as being the result of the algorithm they use which can place the text in one or other category. The workflow we found that works well is then to ensure the messages need-



Near Real-time Processing Challenges

Goal: To find relevant and actionable information in near real-time.



Apply crowdsourcing

Apply machine learning

AIDR

Crowdsourcing + Machine Learning

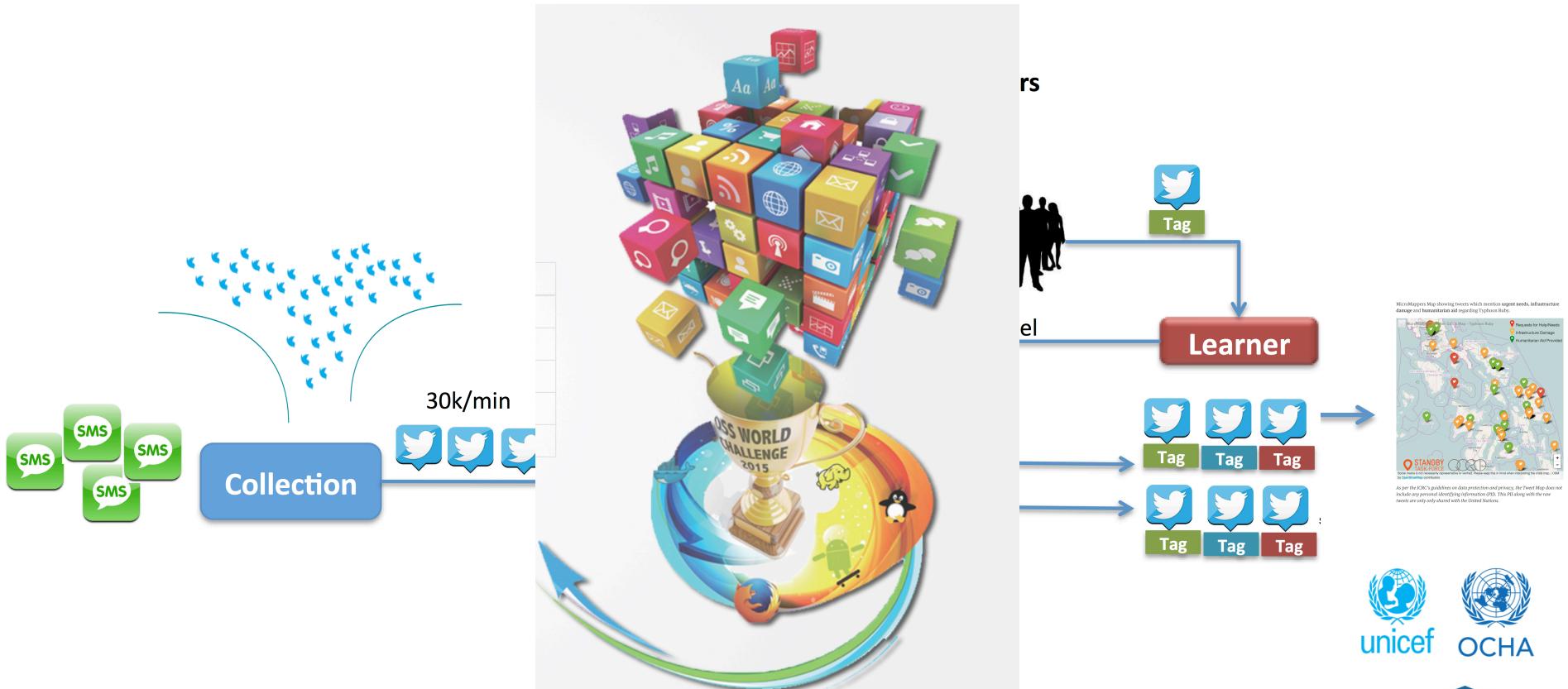


Filter-failure



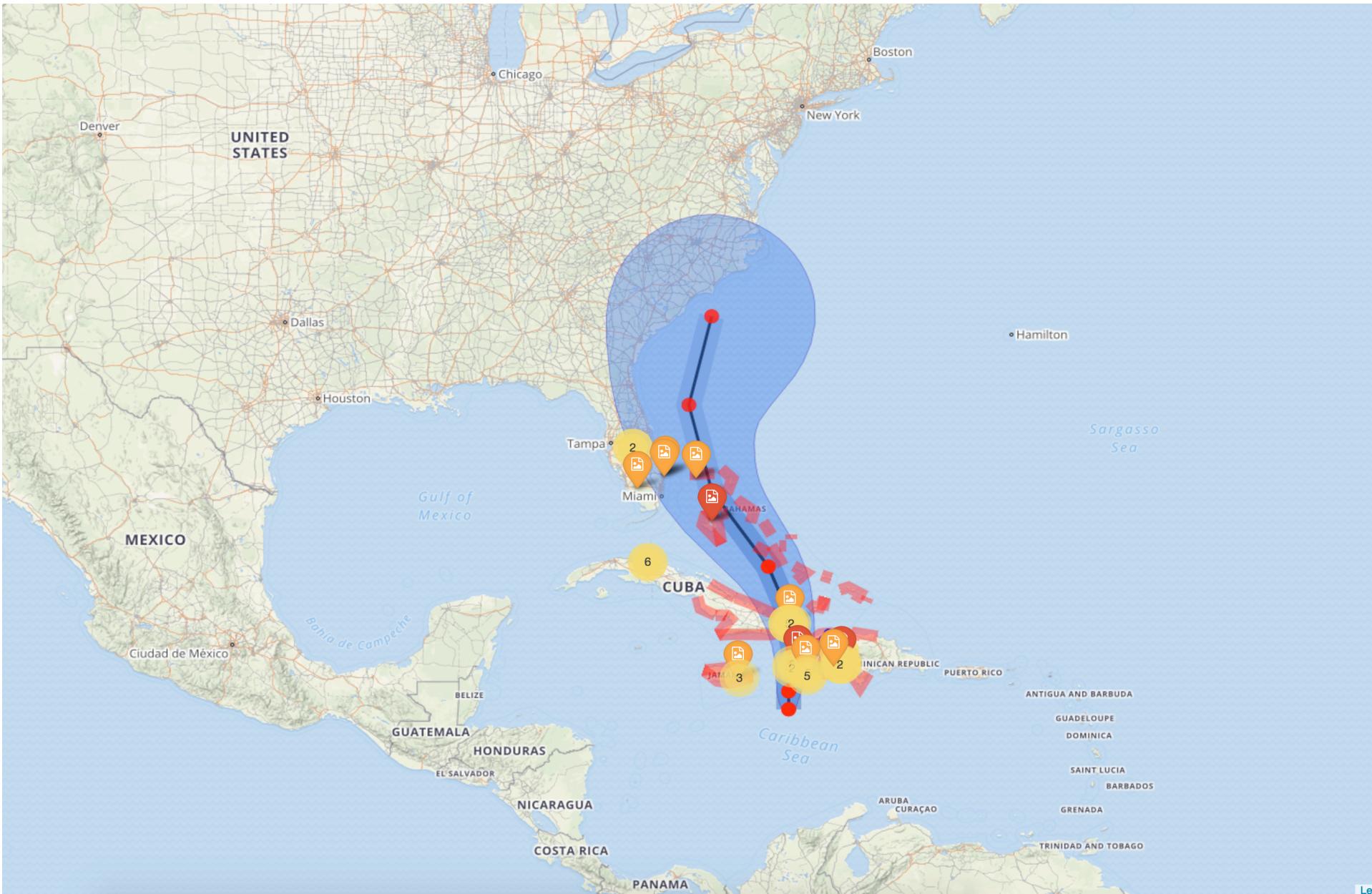
Need human-labeled examples

Real-time Classification of Social Media Data

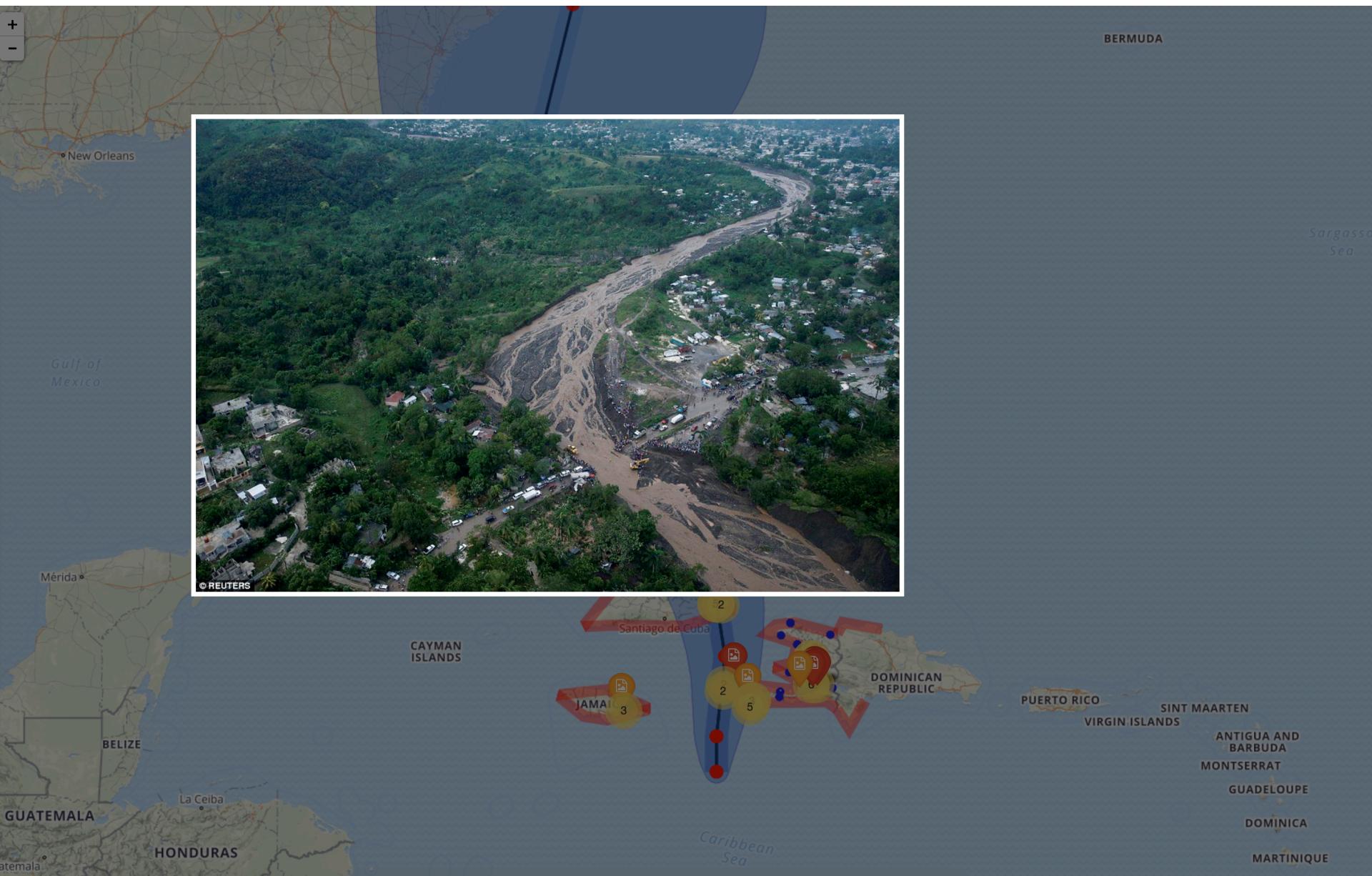


Grand Prize Winner from the Open
Source Software World Challenge 2015

Hurricane Matthew 2016



Hurricane Matthew 2016



“A picture is worth a thousand words.”

Nepal
Earthquake



Ecuador
Earthquake



Hurricane
Matthew



Damage Assessment from Images

How much damage do you see in this picture:



1

None

2

Mild

3

Severe

Relevancy Filtering



Examples of irrelevant images showing cartoons, banners, advertisements, celebrities, etc.

Blurred and unblurred



With and without text



Cropped

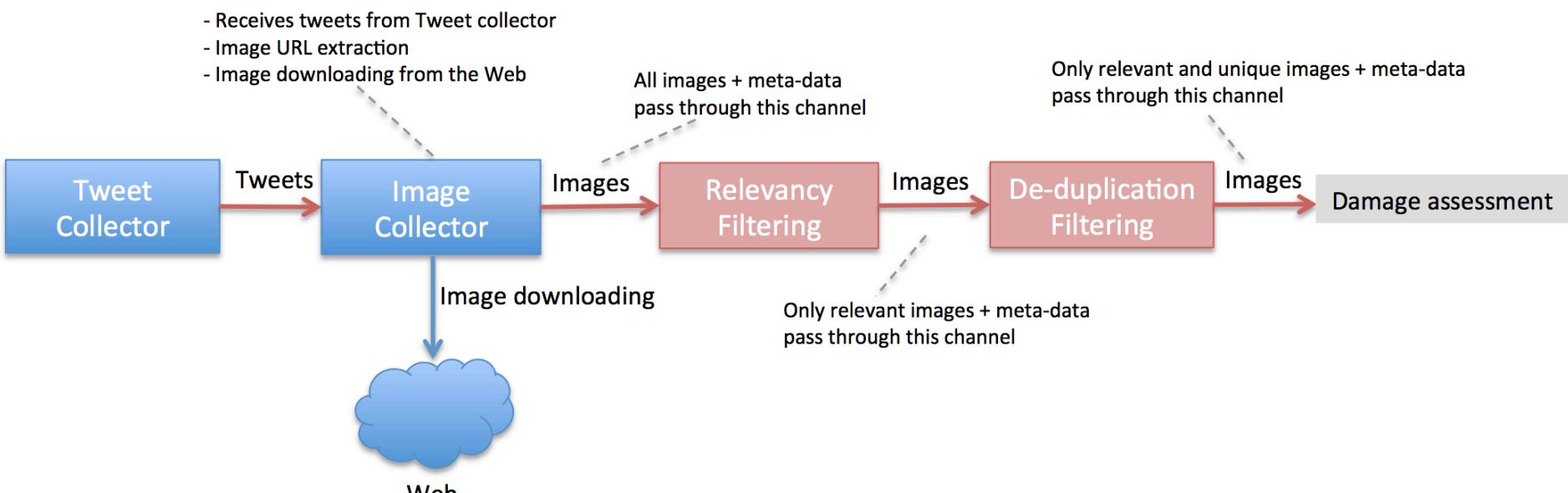


Re-sized

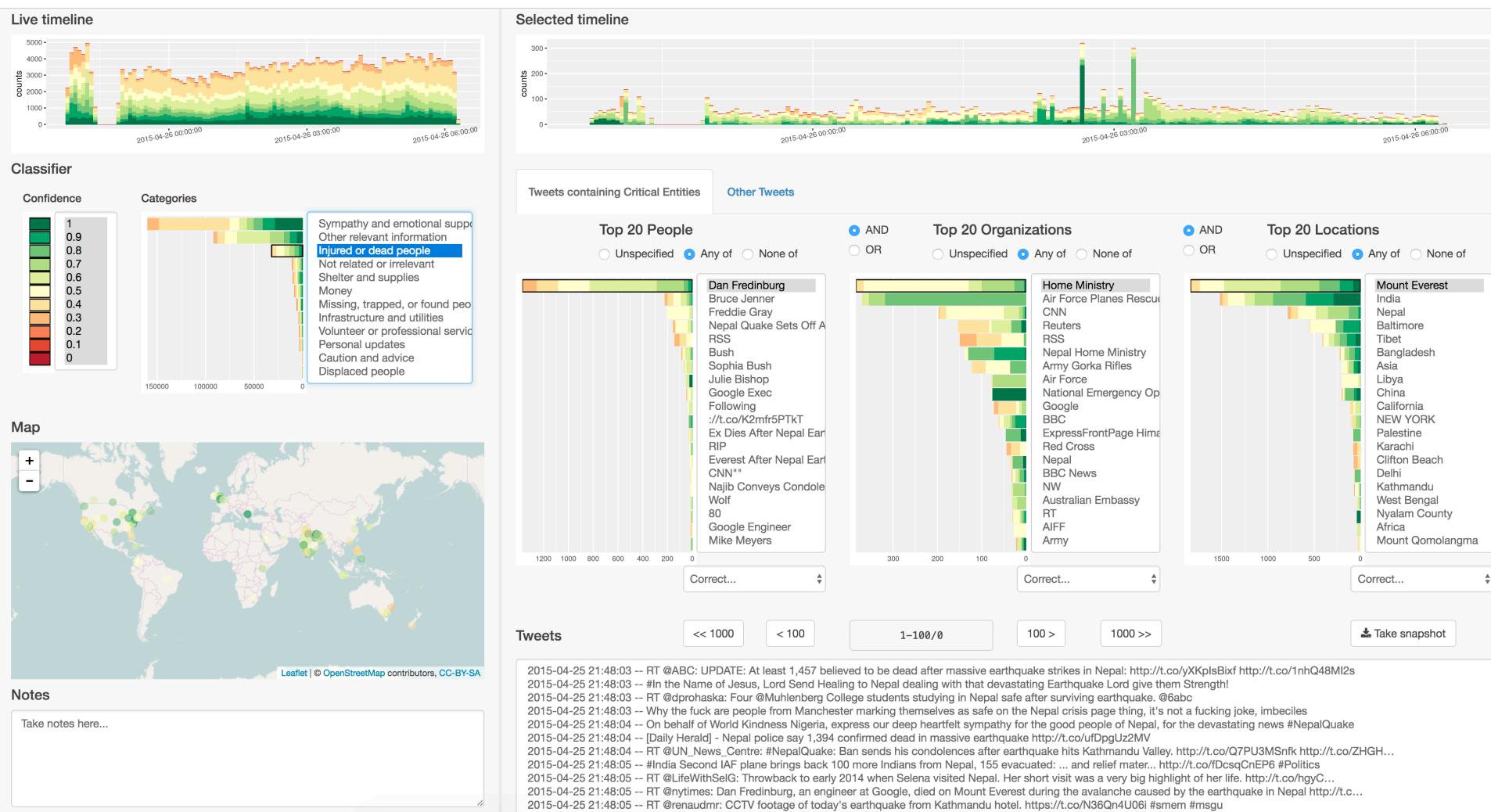


Examples of near-duplicate images

Automatic Image Processing Pipeline



Real-time Situational Awareness Using Interactive Dashboard



Media Coverage

nature Forbes



Mashable

WIRED.CO.UK

New
Scientist

THE WALL STREET JOURNAL.
WSJ

VOA
Voice of America

BROOKINGS

GULF TIMES

Qatar Tribune
First with the news and what's behind it
Nation • Business • Sports • Chill Out

The Peninsula
Local focus. Global vision.
QATAR'S DAILY NEWSPAPER

Potential Applications of AIDR

- Event monitoring and management
- Outbreaks detection, monitoring, and learning about symptoms and treatment
- Insurance claims verification
- Airplanes fault detection in real-time
- Patients to medical center assignments and prioritization

Thank you!

Q & A

Backup slides

Nepal Earthquake 2015

SEVERE
2015-04-26 01:40:05

SEVERE
2015-04-26 01:40:04

MILD
2015-04-26 01:40:03

MILD
2015-04-26 01:40:03

SEVERE
2015-04-26 01:36:43

MILD
2015-04-26 01:36:32

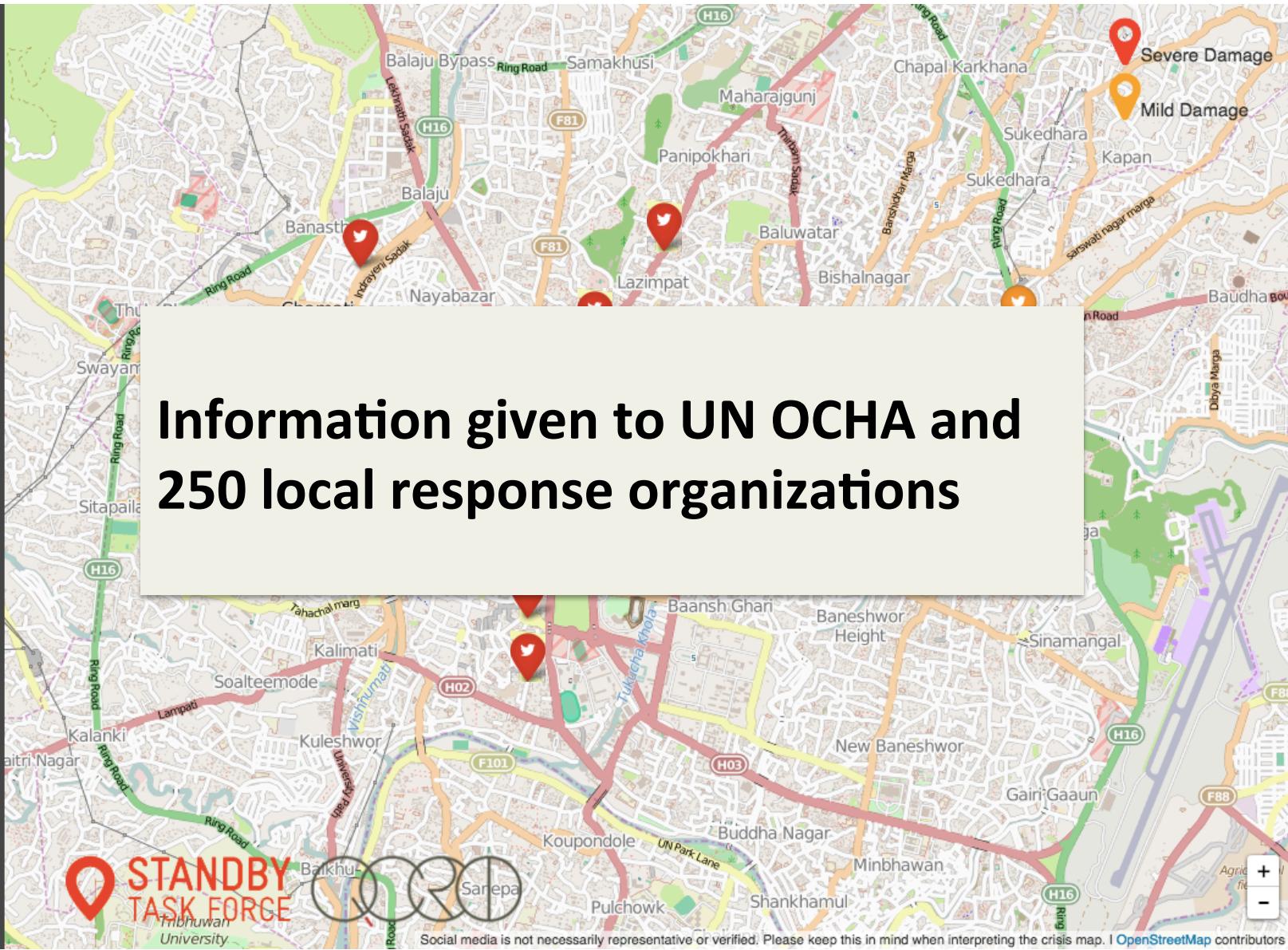
MILD
2015-04-26 01:22:37

MILD
2015-04-26 01:33:09

SEVERE
2015-04-26 01:33:08

MILD
2015-04-26 01:33:08

SEVERE
2015-04-26 01:22:07



Relevancy Filtering



Examples of irrelevant images showing cartoons, banners, advertisements, celebrities, etc.

Task: Build a binary classifier

Approach: Transfer learning

(fine-tune a pre-trained convolutional neural network, e.g., VGG16*)

Performance of the relevancy filtering

AUC	Precision	Recall	F1
0.98	0.99	0.97	0.98

* Simonyan, K. and Zisserman, A. (2014). "Very deep convolutional networks for large-scale image recognition". In: arXiv preprint arXiv:1409.1556

Duplicate Filtering

Blurred and unblurred



With and without text



Cropped



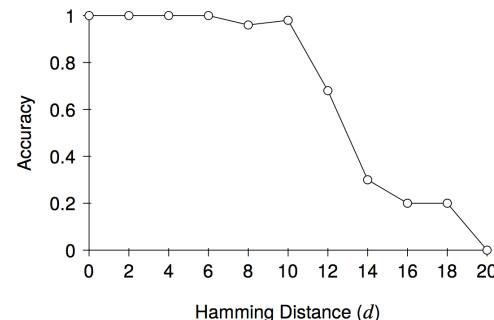
Re-sized



Examples of near-duplicate images

Task: Compute similarity between a pair of images

Approach: Perceptual Hash* + Hamming Distance (w/ threshold)



* Lei, Y. et al. (2011). "Robust image hash in Radon transform domain for authentication".

In: Signal Processing: Image Communication 26.6, pp. 280–288.

Before/After Image Filtering

Number of images that remain in our dataset after each image filtering operation

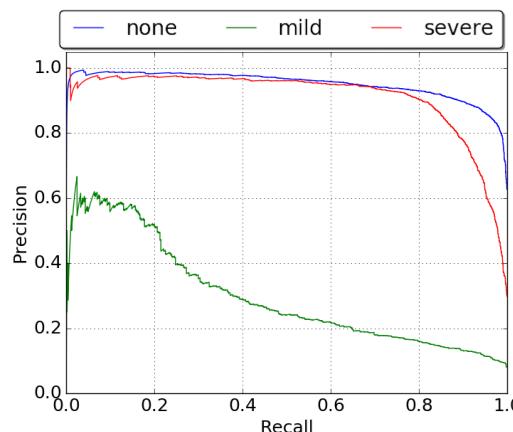
Category	Raw Collection	After Relevancy Filtering	After De-duplication Filtering	
Severe	7,501	~ 2 %	7,289	~ 58 % 3,084
Mild	1,839	~ 2 %	1,675	~ 50 % 844
None	18,186	~ 50 %	9,300	~ 30 % 6,553
All	27,526		18,264	10,481

Assume tagging an image costs \$1, we could have gotten the same job done by paying \$17k less, almost saving 2/3s of the budget!!!

Infrastructure Damage Assessment

- Three-class classification
 - Categories: severe, mild & little-to-none
 - Distinction between categories is ambiguous.
 - Agreement among human annotators is low.
 - in particular for *mild* category
- Fine-tuning a pre-trained CNN (e.g., VGG16)

	AUC	Pre.	Rec.	F1
None	0.95	0.86	0.95	0.91
Mild	0.30	0.55	0.14	0.23
Severe	0.91	0.86	0.85	0.86
Avg.	0.72	0.76	0.65	0.67

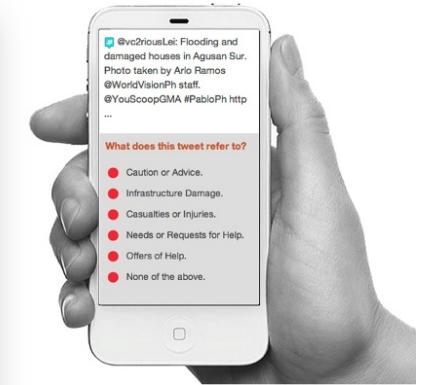


Human Annotation Using MicroMappers

MicroMapper Interface (web clicker)

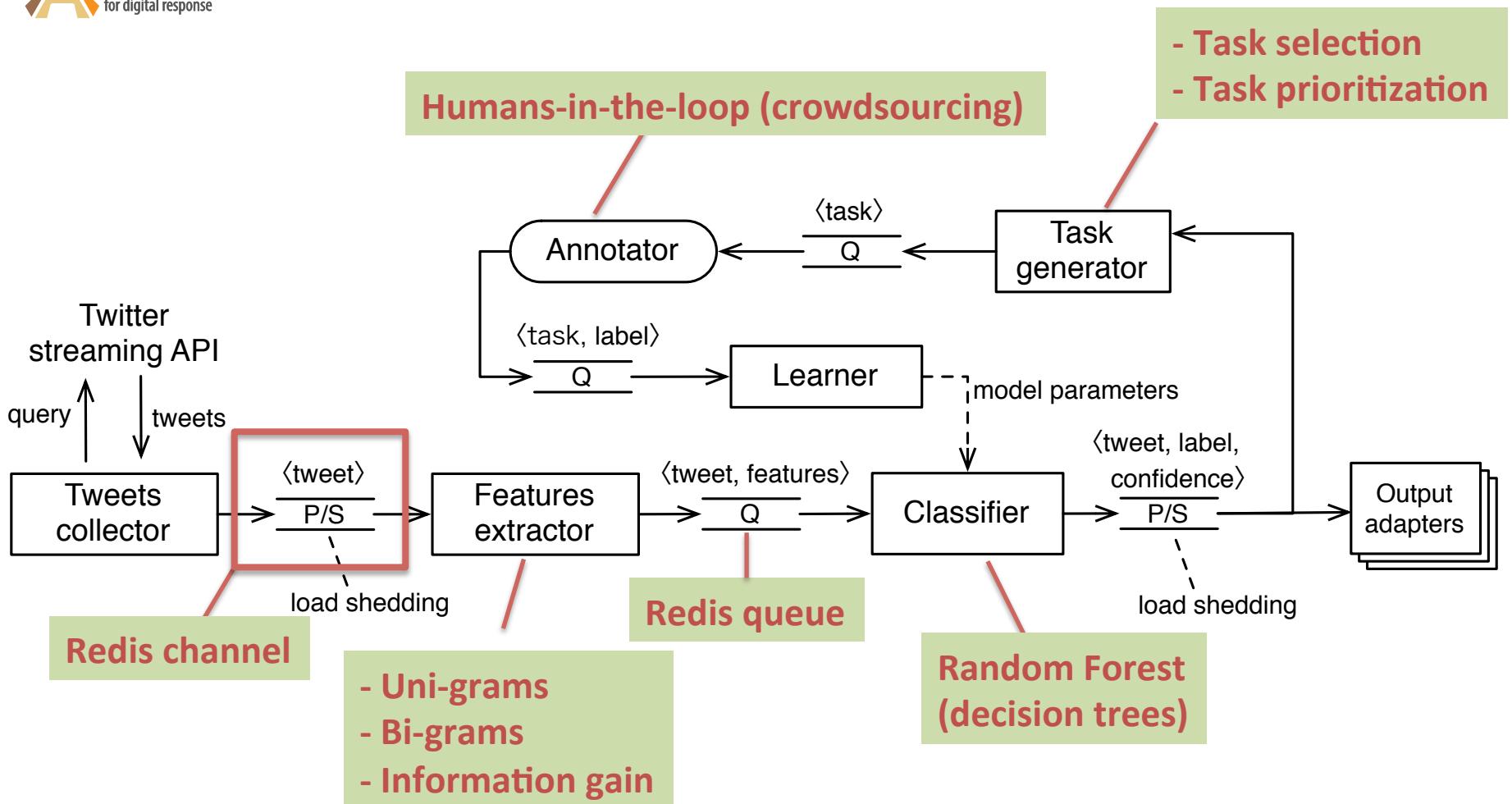
The screenshot shows the MicroMapper web interface. At the top, there's a navigation bar with 'MicroMappers' (highlighted in red), 'Community', 'Applications' (which is the active tab), and 'About'. On the right is a 'Sign in' button. Below the navigation, a project title 'Flu season: Related to flu: MicroMappers' is displayed. A tweet from '@hakaaase' is shown: 'I'm pretty sure I've caught that super bad cold/flu/virus thing here so I'll be doing the same haha gonna scare my momma'. To the right of the tweet, there's a summary box: 'You are working now on task: 22611' and 'You have completed: 25 tweets from 166'. Below this is a 'Tip' section with a link to a 'Tutorial'. At the bottom left is a 'Submit' button. On the far left, there's a list of three radio button options: 'Not related' (selected), 'N/A: does not apply, or cannot judge', and 'Yes related'. At the bottom, it says 'Powered by PyBossa' and 'A project by QCRI'.

Mobile clicker



<http://aidr.qcri.org/>

AIDR Architecture



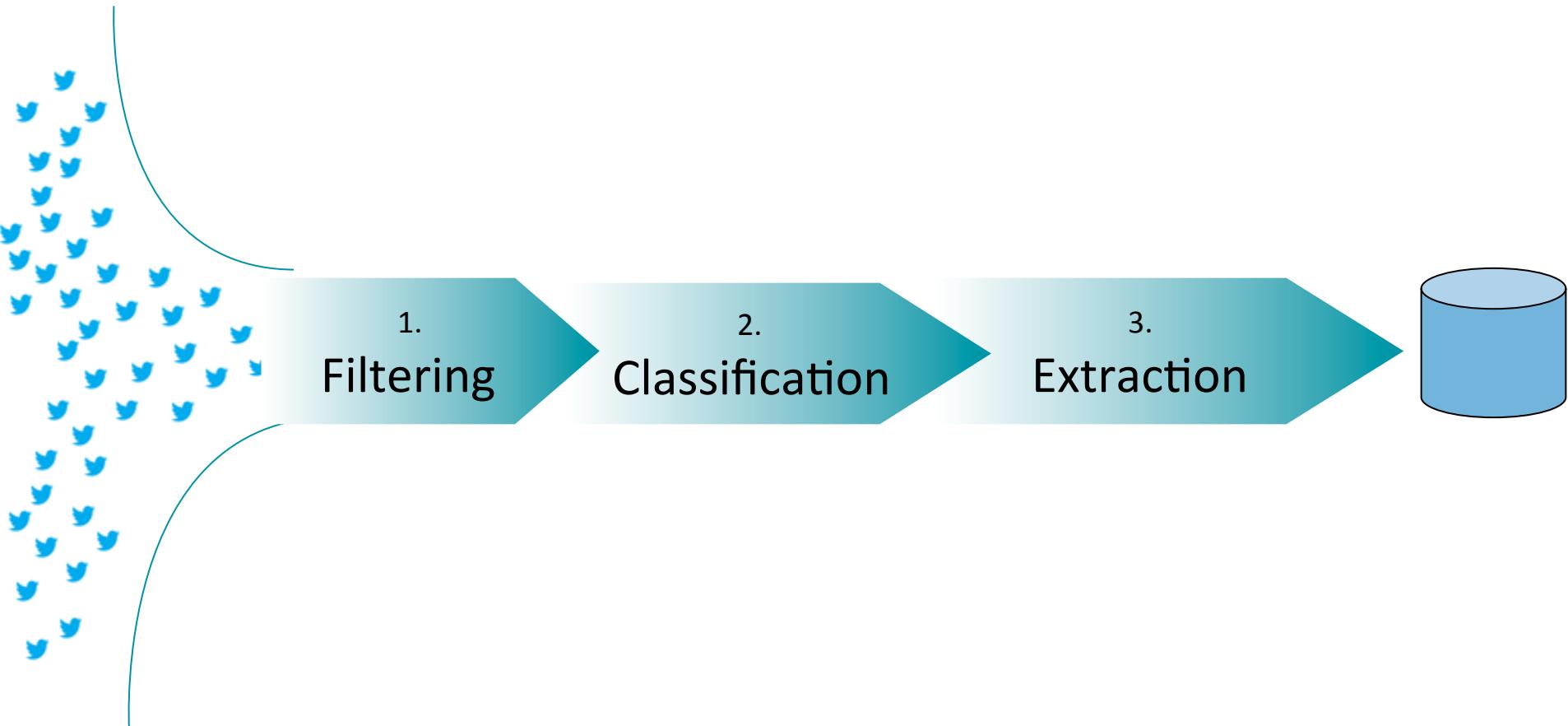
Database: Postgres

Application layer: Java EE, RESTful services, Weka machine learning library

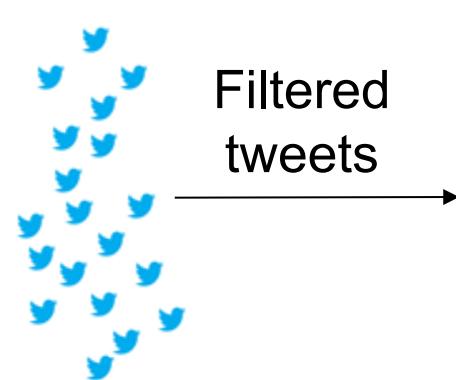
Data flow and control flow: Redis

Front-end: ExtJS (JavaScript library)

Data Classification Approach



2. Classification



Filtered
tweets

	Caution & Advice	Information Sources	Damage & Casualties	Donations	...
Health					...
Water					...
Food					...
Shelter					...
Logistics					...

Sandy Hurricane Twitter Data Analysis



freaking out. home alone. will just watch tv #Sandy #NYC.



@NYGovCuomo orders closing of NYC bridges. Only Staten Island bridges unaffected at this time. Bridges must close by 7pm. #Sandy #NYC.



rt @911buff: public help needed: 2 boys 2 & 4 missing nearly 24 hours after they got separated from their mom when car submerged in si. #sandy #911buff



400 Volunteers are needed for areas that #Sandy destroyed.

Sandy Hurricane Twitter Data Analysis



freaking out. home alone. will just watch tv #Sandy #NYC.

Personal



@NYGovCuomo orders closing of NYC bridges. Only Staten Island bridges unaffected at this time. Bridges must close by 7pm. #Sandy #NYC.



rt @911buff: public help needed: 2 boys 2 & 4 missing nearly 24 hours after they got separated from their mom when car submerged in si. #sandy #911buff



400 Volunteers are needed for areas that #Sandy destroyed.

Informative

Sandy Hurricane Twitter Data Analysis



freaking out. home alone. will just watch tv #Sandy #NYC.

Personal



@NYGovCuomo orders closing of NYC bridges. Only Staten Island bridges unaffected at this time. Bridges must close by 7pm. #Sandy #NYC.

Caution and Advice



rt @911buff: public help needed: 2 boys 2 & 4 missing nearly 24 hours after they got separated from their mom when car submerged in si.
#sandy #911buff

Missing people report



400 Volunteers are needed for areas that #Sandy destroyed.

Donation request

Informative

MERS Outbreak: Twitter Data Analysis

Middle East Respiratory Syndrome (MERS)

Twitter data analysis from: 2014-04-27 to 2014-07-14

Qualitative analysis categories:

Reports of symptoms

Reports of signs or symptoms such as fever, cough or questions

Affected people reports

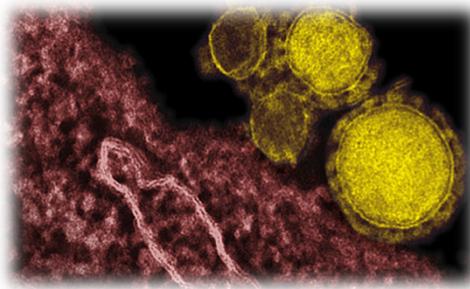
Reports of affected people due to the MERS disease

Death reports

Reports of deaths due to the MERS disease

Prevention questions

Questions or suggestions related to the prevention of disease

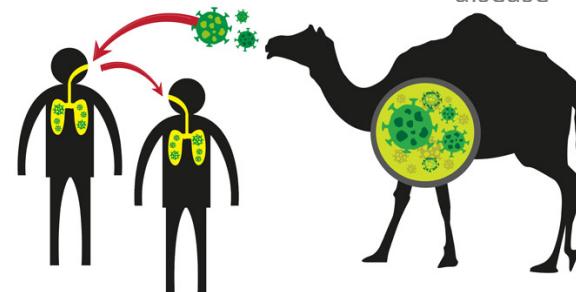


Disease transmission reports

Reports or questions related to the transmission of the disease

Treatment questions

Questions or suggestions regarding the treatment of the disease



Social Media During MERS Outbreak



RT @abettel: Two workers at **FL** hospital exposed to a patient with Middle East Respiratory Syndrome are **showing flu-like symptoms**

Signs and symptoms



Coronavirus symptoms include: **fever, coughing, shortness of breath, congestion in the nose and throat, and in some cases diarrhea.** MERS

Signs and symptoms



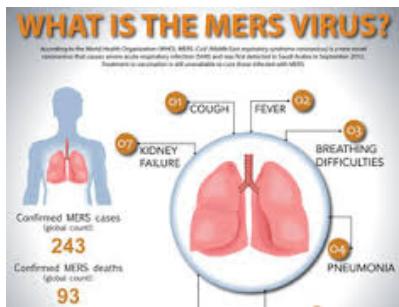
#MERS is a relatively new respiratory illness, spread b/w people in close contact. **Symptoms are fever, cough, & shortness of breath.**

Signs and symptoms



Saudi Arabia finds another **32 MERS cases** as disease spreads: RIYADH (Reuters) - Saudi Arabia said on Thursday ... <http://t.co/cPhm0uTRCo>

Affected individuals



Social Media During MERS Outbreak



First Case of Deadly Middle Eastern Virus Found in U.S.: The Centers for Disease Control has confirmed that a case of the deadly Midd...

Affected individuals



Third Case of MERS Confirmed in the U.S.: The U.S. Centers for Disease Control and Prevention confirmed on Sat... <http://t.co/Sb8PMyxVUn>

Affected individuals



No clear transmission link btwn camels and humans for MERS. 94% Egyptian camels seropositive but no human cases yet. Hmm #asm2014

Transmission



Saudi health authorities announced on Monday that the **death toll from the MERS coronavirus has reached 115** since the respiratory disease ...

Death reports



Twitter Breaks Events Faster

Hudson Plane Crash



First report



Janis Krums
@jkrums



 Follow

<http://twitpic.com/135xa> - There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy.

RETWEETS
244

LIKES
878



9:36 PM - 15 Jan 2009



Westgate Mall Attack

Breaks the story 33 minutes before local TV



I come from Somalia.
@geekinthejungle



 Follow

gun shots in westlands? wtf??

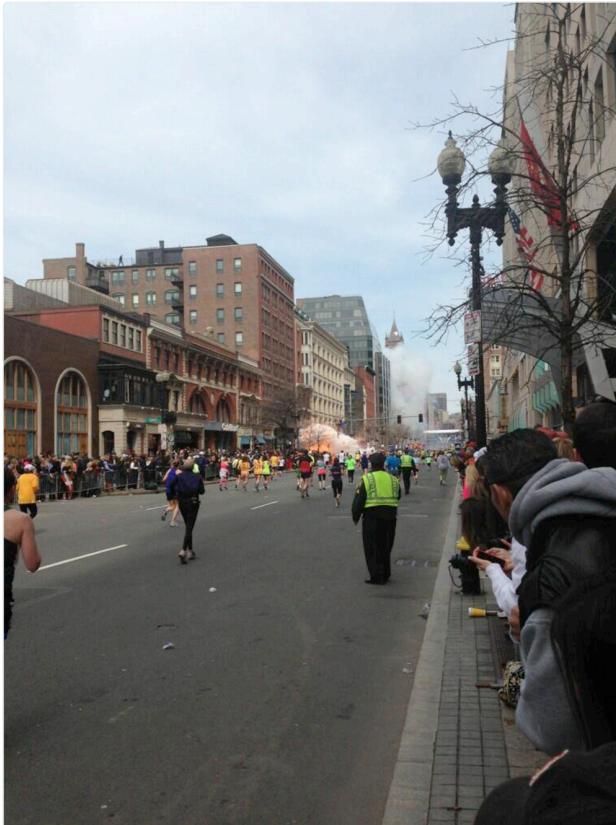
 Reply  Retweet  Favorite  Buffer  Storify  More 

12:38 PM - 21 Sep 13

Twitter Breaks Events Faster

Boston Bombing

First report on Twitter



Dan Lampariello
@WBBJ7Dan

Explosion at copy

9:50 PM - 15 Apr 2013

↳ 1,943 ⚡ 153 ❤ 187

After 1 minute

Pure Applesauce
@stackii

[Follow](#)

An explosion just went off in downtown Boston. Spectators fleeing the #bostonmarathon course.

9:51 PM - 15 Apr 2013

↳ 79 ⚡ 11

Pure Applesauce
@stackii

[Follow](#)

I think a bomb just went off in Boston. Can't tell. Can smell smoke. Emergency vehicles everywhere.

instagram.com/p/Yltk1gTIUJ/

9:53 PM - 15 Apr 2013

↳ 308 ⚡ 21

After 2 minutes

Fox Sports 1380 KRKO
@KRKO1380

[Follow](#)

BREAKING: Per our man on the ground at the Boston Marathon, @toobladogs, there was an explosion. More to follow.

9:52 PM - 15 Apr 2013

↳ 153 ⚡ 3

Data and Opportunities



2013 Pakistan Earthquake
September 28 at 07:34 UTC



Disease outbreaks



2010 Haiti Earthquake
January 12 at 21:53 UTC

Availability of Immense Data:

 Twitter  @twitter

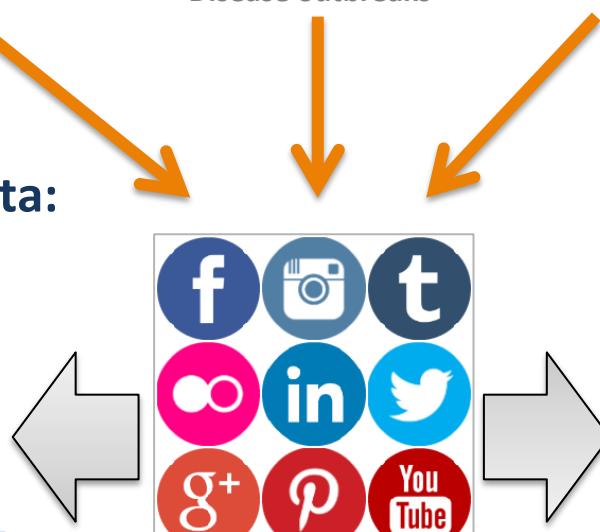
Follow

People sent more than 20 million Tweets about the storm between Oct 27 & Nov 1. Terms tracked: "sandy", "hurricane", #sandy, #hurricane.

6:46 PM - 2 Nov 2012

619 RETWEETS 199 FAVORITES

Around 16 thousands tweets per minute were posted during the hurricane Sandy in the US.

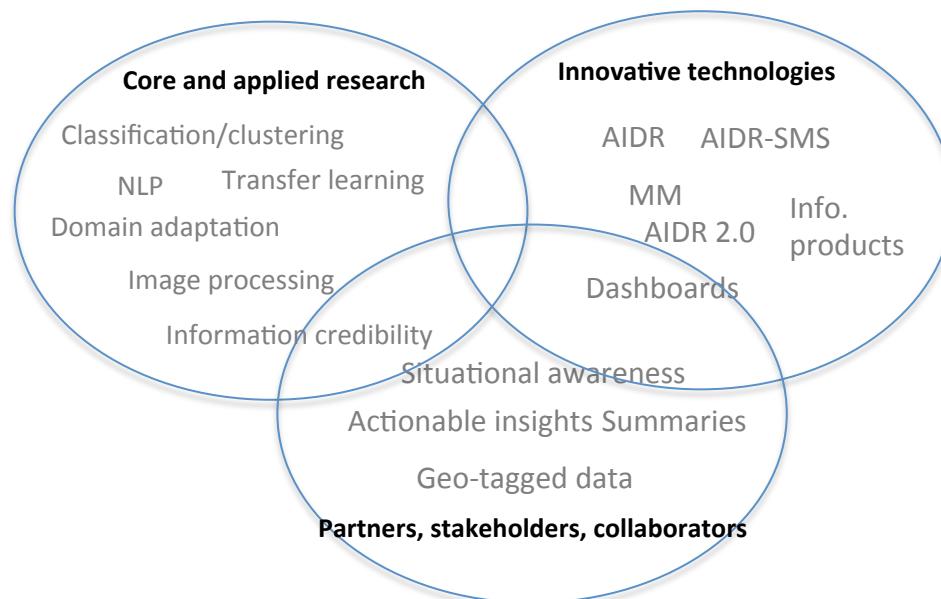


Opportunities:

- Early warning and event detection
- Situational awareness
- Actionable information extraction
- Rapid response
- Effective communications

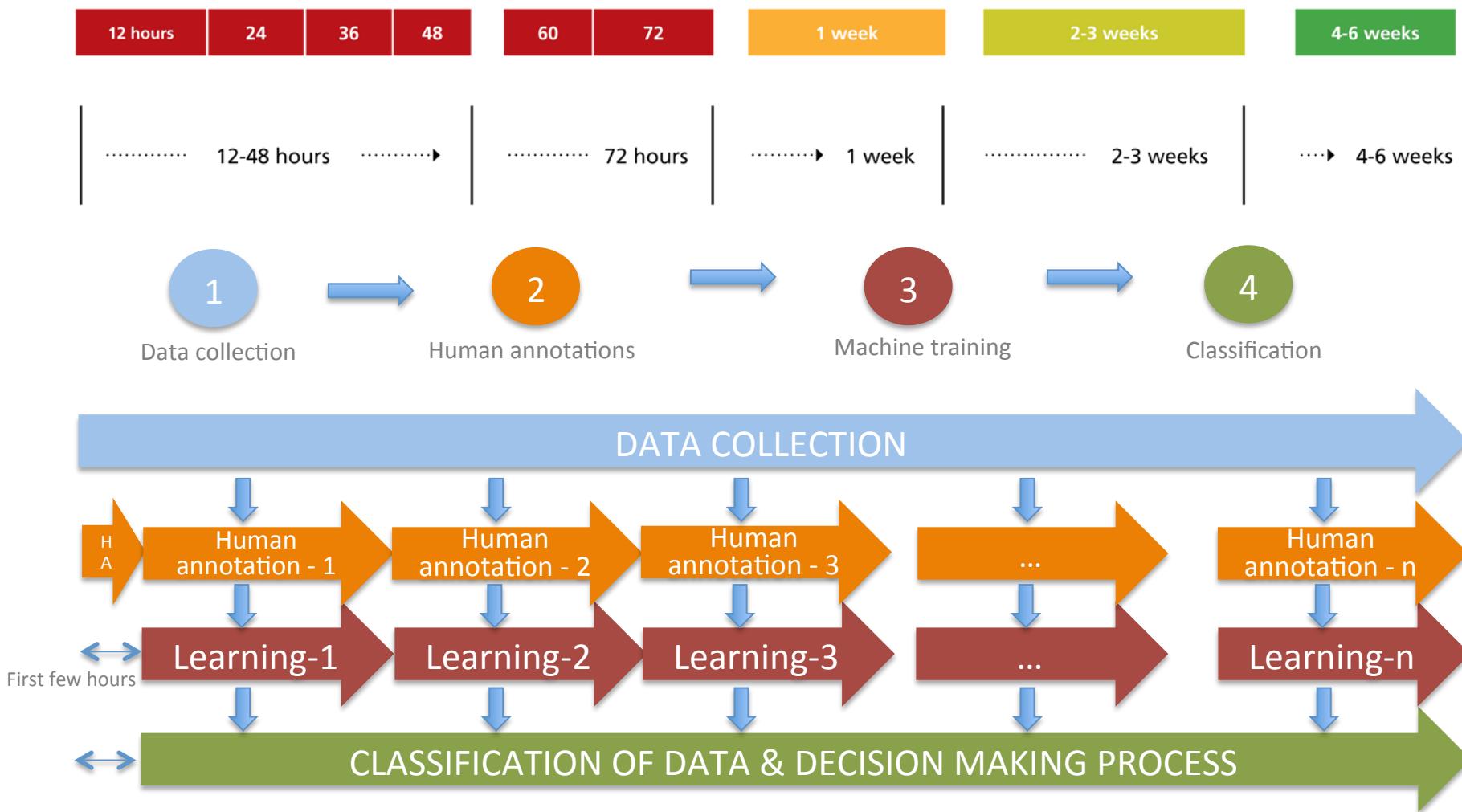
Vision and Focus of AIDR

*"To help **policymakers, NGOs, affected communities, and scholars** improve the effectiveness of humanitarian strategies such as **preparedness, mitigation, and response** during natural disasters, crises and emergencies."*



Near Real-time Processing

ONLINE APPROACH



Systems for Crisis-Relevant Data Processing

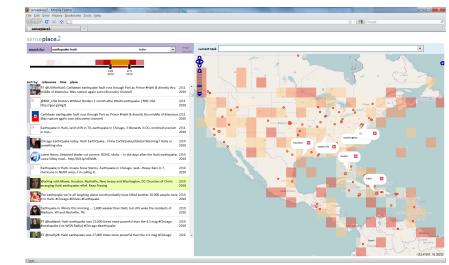
Twitris [Purohit and Sheth 2013]

Twitter; semantic enrichment, classify automatically, geotag



SensePlace2 [MacEachren et al. 2011]

Twitter; geotag, visualize heat-maps based on geotags



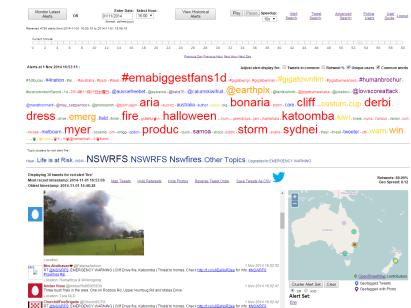
EMERSE Enhanced Messaging for the Emergency Response Sector [Caragea et al. 2011]

Twitter and SMS; machine-translate, classify automatically, alerts



ESA Emergency Situation Awareness [Yin et al. 2012; Power et al. 2014]

Twitter; detect bursts, classify, cluster, geotag



Systems for Crisis-Relevant Data Processing

Twitcident [Abel et al. 2012]

Twitter and TwitPic; semantic enrichment, classify

The screenshot shows a search results page for "Texas Fire". At the top, it says "Started on Sep. 1, 2011 12:00:00 AM" and "Current selection: 22,968 tweets users". Below this is a list of tweets from various users like "objective21", "TeleSetLamLusk", "Deposito", and "traveler7200". On the right side, there are filters for "Search in Texas Fire tweets", "Filter tweets", and "Remove filters".

CrisisTracker [Rogstadius et al. 2013]

Twitter; cluster, annotate manually

The screenshot shows the Crisis Tracker interface with a map of Syria and Lebanon. The map has several red dots indicating event locations. On the left, there are filters for "WHAT" (Civilian involvement, Military involvement, Crime, Violence, Missing people, Damaged infrastructure), "WHERE" (Natural hazard, Political event, Summary report, Available resource, Request/Need, Warning/Risk/Danger, High impact event), "WHO" (Enter name), and "WHEN" (From, To). Below the map, there is a list of events with details like date, time, title, and tags.

Tweedr [Ashktorab et al. 2014]

Twitter; classify automatically, extract information, geotag

AIDR: Artificial Intelligence for Disaster Response [Imran et al. 2014a]

Twitter; annotate manually, classify automatically

The screenshot shows the AIDR interface with a collection titled "Global HIV-Related". It lists tweets from users like "Global HIV-Related", "Zika Virus in Brazil", and "EMSC Landslides by KW". Each tweet has a "Collect" and "Classify" button next to it. The interface also includes a sidebar for "My Collections" and a "Create Collection" button.