THE UNIVERSITY OF TEXAS AT EL PASO

Department of
Computer Science

Sept. 19, 2016

Dear Participants,

We are seeking your input on new functionality we are developing to support intelligence analysts in their work.

DARPA's Lorelei program aims to support analysts faced with large amounts of data, including data from languages they do not know. The data might come from radio news, social media, or other sources. As part of Lorelei, the UTEP-SRI research team has developed techniques for automatically inferring the speaker's attitude to what he or she is describing, that is, the "stance" of the speaker.

We are now soliciting feedback on how useful this would be, and how to make it more useful. We are soliciting input from two Army units, one Air Force unit, and a DoD agency. We have prepared scenarios of where stance might be useful, and mock-ups of an interface that an analyst might see. We would like you to go through these, answer some questions, and provide any comments or suggestions. We expect that this will take just over an hour.

We will use your feedback to help decide which aspects of stance to prioritize for improving performance. We will also send an anonymized summary to DARPA, to help them decide whether to keep stance as part of the Lorelei program. There is no need to identify yourself to us. All questions are optional.

We apologize for not being able to meet with you ourselves. However my colleague Melissa Caraway should be able to answer most questions. Please give the forms and feedback to him. We also welcome direct contact; you can reach us at nigel@utep.edu.

Thank you!

Nigel G. Ward

Nigel G. Ward, Professor, PI

Olac Fuentes

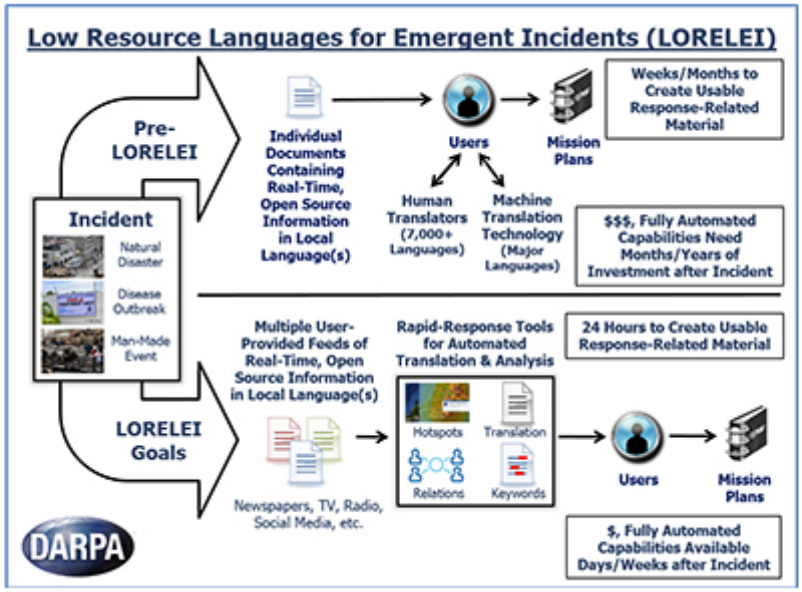
Olac Fuentes, Associate Professor

El Paso, Texas
79968-0518
(915) 747-5480

LORELEI Imagines Rapid Automated Language Toolkit

Low-cost approach to extracting critical concepts from public information sources in unfamiliar languages would support disaster relief and other quick-response missions

OUTREACH@DARPA.MIL
10/8/2015



Understanding local languages is essential for effective situational awareness in military operations, and particularly in humanitarian assistance and disaster relief efforts that require immediate and close coordination with local communities. With more than 7,000 languages spoken worldwide, however, the U.S. military frequently encounters languages for which translators are rare and no automated translation capabilities exist. DARPA's Low Resource Languages for Emergent Incidents (LORELEI) program aims to change this state of affairs by providing real-time essential information in any language to support emergent missions such as humanitarian assistance/disaster relief, peacekeeping and infectious disease response. The program recently awarded Phase 1 contracts to 13 organizations.

"The global diversity of languages makes it virtually impossible to ensure that U.S. personnel will be able to understand the situation on the ground when they go into new environments," said Boyan Onyshkevych, DARPA program manager. "Through LORELEI, we envision a system that could quickly pick out key information—things such as names, events, sentiment and relationships—from public news and social media sources in any language, based on the system's understanding of other languages. The goal is to provide immediate, evolving

situational awareness that helps decision makers assess and respond as intelligently as possible to dynamic, difficult situations."

The conventional system of developing automated language technology—which requires years of effort and tens of millions of dollars to manually translate, transcribe and annotate individual words and phrases for each language—is adequate for languages in widespread use or in high demand. It is neither flexible enough to meet constantly changing language needs, however, nor specialized enough to account for the specific communication challenges involved in military-level emergency response.

LORELEI seeks to dramatically advance computational linguistics and human language technology to identify the elements that different languages have in common, and use that knowledge to enable rapid, low-cost development of automated language capabilities. The program would apply these automated capabilities via an easy-to-use interface that would assimilate, integrate and analyze real-time incident data in the local language(s). The envisioned system would provide useful response-related material as quickly as 24 hours after an incident occurs and fully automated language capabilities within days or weeks after that.

While LORELEI technologies could include partially or fully automated speech recognition and/or machine translation, the program does not primarily seek to comprehensively translate low-resource languages into English. Instead, LORELEI would provide situational awareness by identifying and correlating elements of information in foreign-language and English sources. LORELEI technology would be applicable to any incident where a sudden need emerges for assimilation of information by U.S. government entities about a region of the world where low-resource languages are frequently used.

"Our goal with LORELEI isn't rote translation based on libraries, but instead to provide idiomatic understanding of language as a whole, and specifically disaster-response vocabulary, to improve cooperation and speed response to dangerous situations worldwide," Onyshkevych said.

DARPA has awarded Phase 1 contracts for LORELEI to the following organizations:

Appen

Carnegie Mellon University

Columbia University

Johns Hopkins University

Next Century Corporation

Raytheon BBN

University of Illinois Urbana-Champaign

University of Massachusetts

University of Pennsylvania

University of Pennsylvania Linguistic Data Consortium

University of Texas El Paso

University of Washington

Stance in Lorelei: Use-Case Scenarios

Background: Stance relates to the speaker's attitude to what he or she is describing. Although an intrinsic aspect of news broadcasts, tweets etc., stance has not previously been used in information retrieval or filtering. While related in some ways to topic and emotion, it provides a novel view.

We are developing automatic stance detection methods. Although these will never be 100% accurate, with large-volume data they should be adequate to reveal patterns.

The scenarios below, and the associated rough mock-ups, illustrate how and why we think stance could help analysts. We seek input on which stances are useful and how to provide this information in an interface.

Scenario A. During a flood, the Lorelei heatmap shows relevant talk across Valleys C and D, with satellite imagery also showing both hard-hit (Mockup A1, shading indicating desity of conversations mentioning "flooding," "water," etc.) The mission planner pulls up the frame containing stance-related sliders (Mockup A2). Dragging these to explore, he finds that Valley D mostly disappears from view after adjusting to exclude talk with a stance of *typical* (Mockup A3). He asks his assistant to first start planning an intervention for Valley C. Later he discovers that in Valley D had a long history of flooding, and there the well-drilled population was calmly evacuating, but C was unprepared.

Scenario B. In a famine situation, the analyst is looking for information on where the need is. Lorelei's heatmap shows blobs in the northern hills and and also in one southern valley (Mockup B1). Knowing that people will talk about a famine even if it is not affecting them directly, he varies the *locally relevant* slider, expecting to see an effect on the blob shapes and densities (B2), but the blobs don't change. This makes him suspect that the initial heatmap is untrustworthy, so he switches off the "famine-related" topic filter and then drags the stance sliders to filter for *bad situation, urgent/immediate, factual,* and *locally relevant*. This time the display shows that the need is only in the northern hills (B3). Later he discovers that the southern dialect uses, the word for "rice" is slang for a bribe, explaining the spurious result.

Scenario C. After a cattle disease affecting a wide area, the analyst looks at a Lorelei heat map, trying to get an overview of the effects and the human terrain. While experimenting with displaying averages on various stances, he discovers that the distribution of feelings mostly matches what he'd expect from the location of the disease, plus the general tendency of the population in that part of Africa to create conspiracy theories to explain any bad luck. But he also notes that conversations from one city, just outside the affected area, have a different mood: low on stance *deplorable* (Mockup C1). He pulls out his shoebox of news clippings on that city, and finds an article about a mega-mosque with an active charity arm and youth league. He then shifts to a social-organization view, and chooses for display only material by and about the mega-mosque and its iman (C2, green). Selecting as the comparison group recent tweets from the rest of the province, and looking at the stance distributions side by side, the difference in tone is clear on stances *praiseworthy action,* and also on *factual* (C2). Hoping that this means that the mega-mosque community has a "can do" attitude and good organizational skills, he googles it. He finds no web presence except a blog with one English post, a

crudely-worded plea for the international community to "let us help you help our brothers." In his report to the local commander he notes the potential opportunity. Later he finds that agreement was reached and the mosque's youth league traveled in with the convoy to help distribute supplies in the affected area.

Scenario D. After a major wildfire, ethnic group A petitions the government to protect them from attacks by ethnic group B. To determine how this might complicate aid operations, the analyst examines material originating from group A on the topic of group B. From past experience, he expects to see stances high in low in *subjectivity,* and high in *urgency/immediate action needed,* but finds these mostly lacking in this case. He then opens the timeline view and examines the stance profiles of material originating from group A over the past year. Detecting an increase in volume, but no big changes in tone before or after the wildfire, he decides that this is not a high-priority concern. Later he learns that Group A had been exploiting the attention about the wildfire to try to enlist public opinion against Group B.

Scenario E. After a storm hits power lines around cities E and F, there is lots of messaging from both relating to electricity. Attempting to determine where the situation is most severe, the analyst applies the *relevant to a large group* stance filter. Finding that F almost disappears from the heatmap, he gives priority to translating messages from E. Later he finds that most of the messaging from city F was about a bizarre but heartwarming event involving a little girl, a downed power line, and the family pet.

Scenario F. Trying to assess the scale of a marketplace bombing in City G, the analyst also finds a lot of messaging on the topic of violence in nearby City H. Applying the *urgent/immediate*, *factual information*, and *new information* filters, he decides that the talk in H is less relevant and decides to focus on G. Later he finds that the messaging in H was about a bombing they had experienced years ago, and opinions about the way it was handled by the authorities, including some politicians now up for re-election.

Scenario G. After an earthquake, the Lorelei heatmap shows two dense clusters of talk about the topic, one in Province A and one in Suburb B. The analyst applies the *locally relevant* filter to cull out talk that has a second-hand/non-local stance, and infers that Province A is the actual location with need. Later he finds out that suburb B was largely populated by the ethnic group from Province A, and the talk there was mostly about concern for relatives in the province.

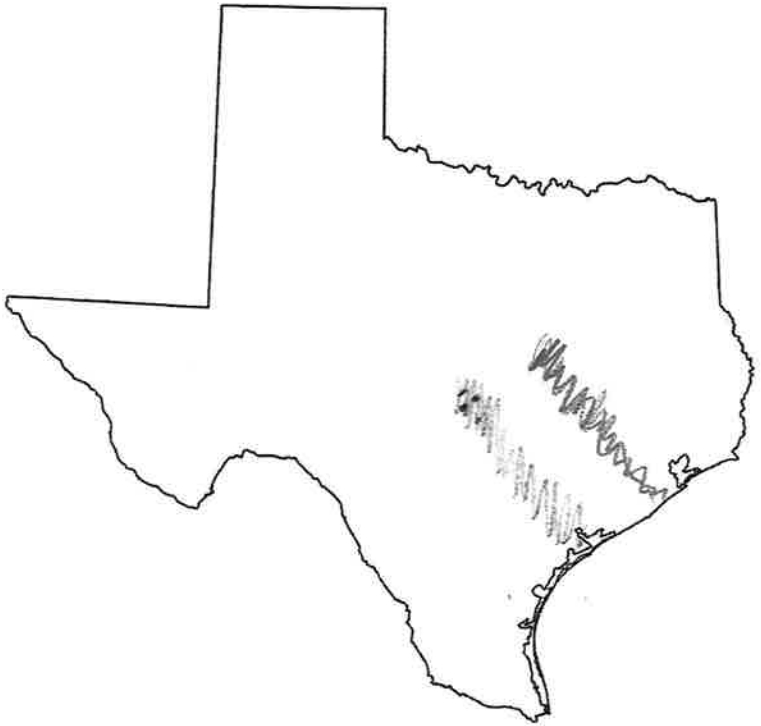
Scenario H. Seismography suggests, and imagery confirms, a volcanic eruption in Province N. The analyst monitors chatter originating from that area (Mockup H1), and locates the hotspot of conversations about this. Looking at the stance averages, he finds the expected increase in volume of discussion and in negative affect. Viewing the stance profiles on a timeline, however, he notes that the *urgent/immediate* stance is lacking, and that *deplorable* spiked three hours after the event (H2, gray line). Feeling that this is oddly early, as people usually don't start blaming the authorities for things until a few days later, he suspects this may not be a true disaster. To confirm his suspicion, he overlays the typical stance averages as they evolved over type for other disasters (H2, blue line). Based on this, he decides not to wake up the decision-maker. Later he learns that the eruption had never posed a real threat, and the angry chatter was by tour guides and hotel owners upset about the government's decision to suspend tourist helicopter flights over the volcano.

A1

Keyword Density for: Flooding

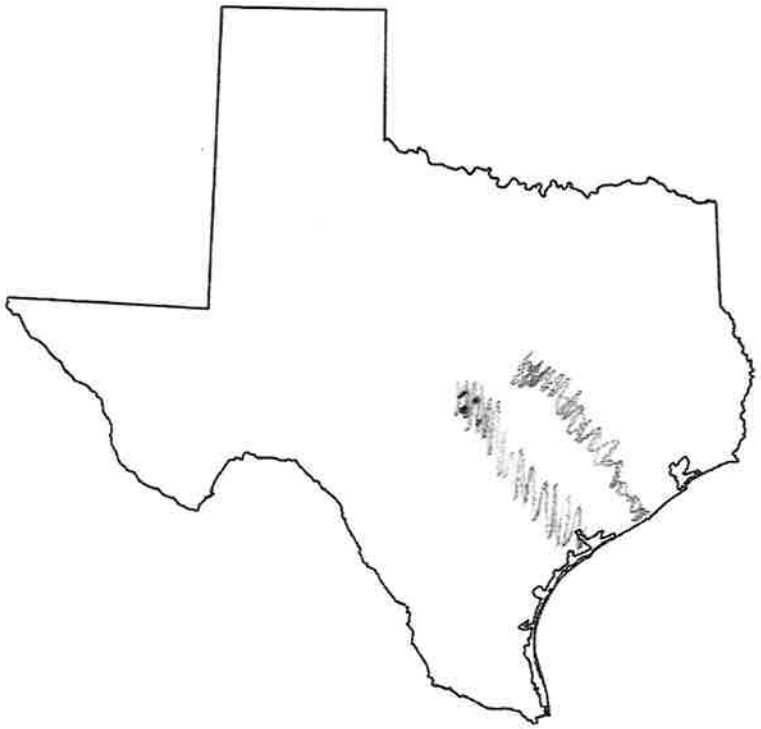
Filtering

- Location
- Time
- Issue
- Emotion
- Stance



A2

Keyword Density for: Flooding

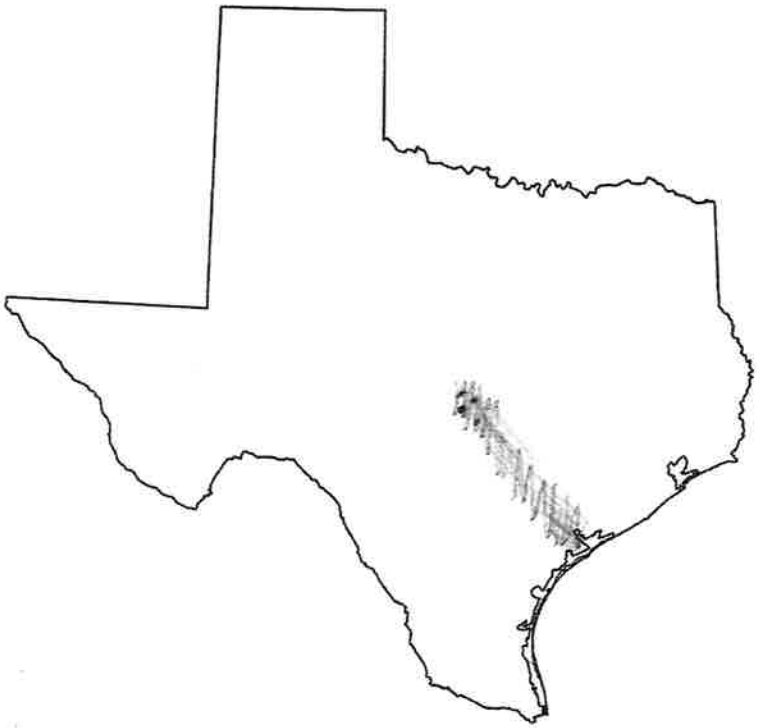


Stance Filtering

Bad	Good
Deplorable	Praiseworthy
Subjective	Factual
Typical	Unusual
Distant	Local
Just Talk	Call to Action
Old	New
Idiosyncratic	Large-Group

A3

Keyword Density for: Flooding



Stance Filtering

Bad — Good

Deplorable — Praiseworthy

Subjective — Factual

Typical — Unusual

Distant — Local

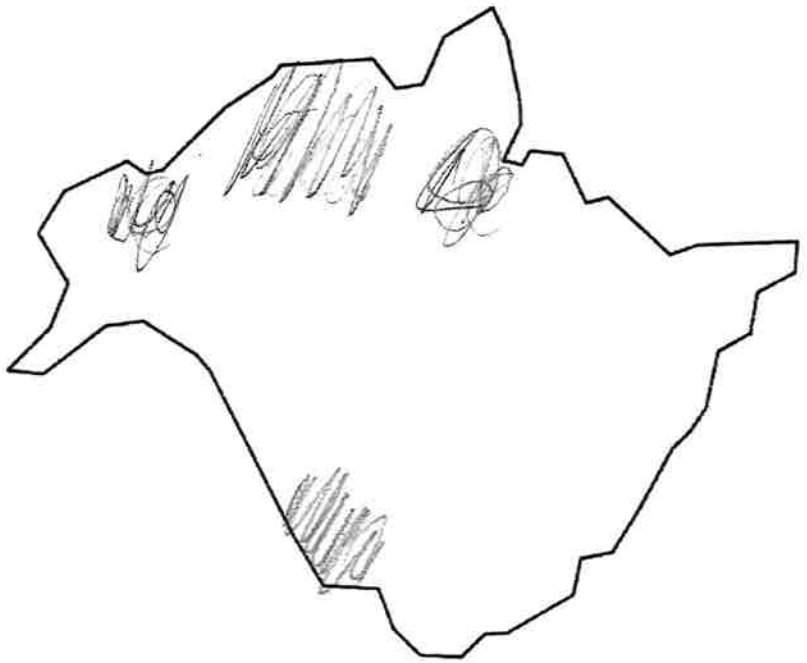
Just Talk — Call to Action

Old — New

Idiosyncratic — Large-Group

B1

Keyword Density for: Famine

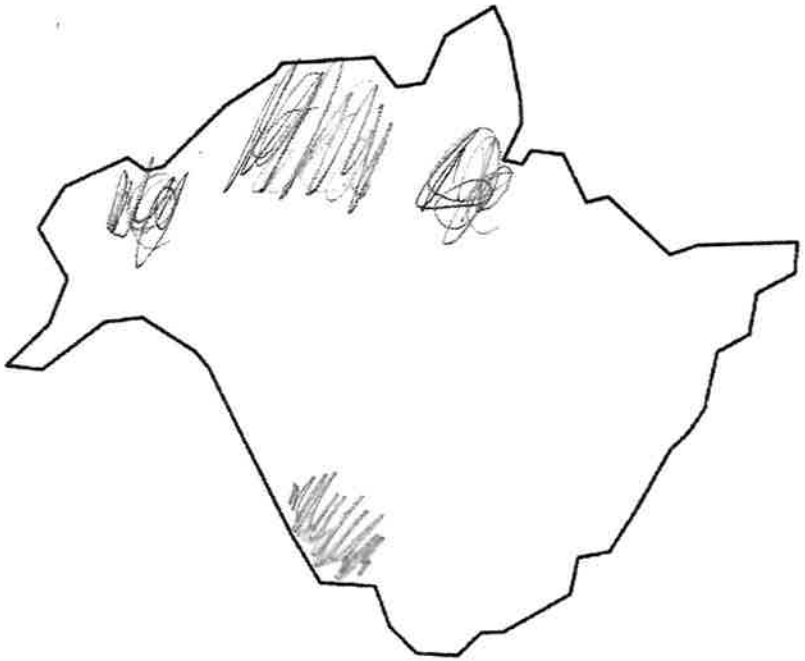


Stance Filtering

Bad	Good
Deplorable	Praiseworthy
Subjective	Factual
Typical	Unusual
Distant	Local
Just Talk	Call to Action
Old	New
Idiosyncratic	Large-Group

B2

Keyword Density for: Famine



Stance Filtering

Bad — Good

Deplorable — Praiseworthy

Subjective — Factual

Typical — Unusual

Distant — Local

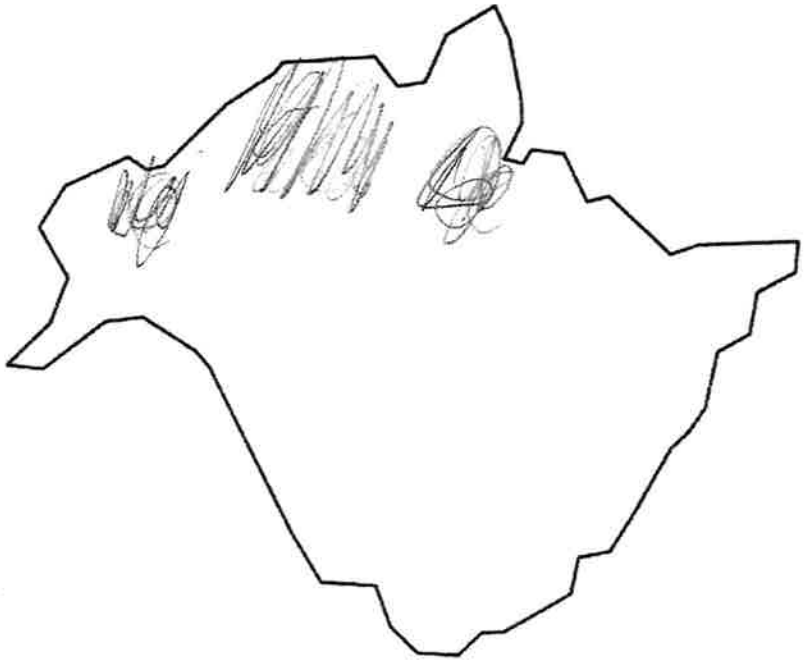
Just Talk — Call to Action

Old — New

Idiosyncratic — Large-Group

B3

Keyword Density for: any



Stance Filtering

Bad — Good

Deplorable — Praiseworthy

Subjective — Factual

Typical — Unusual

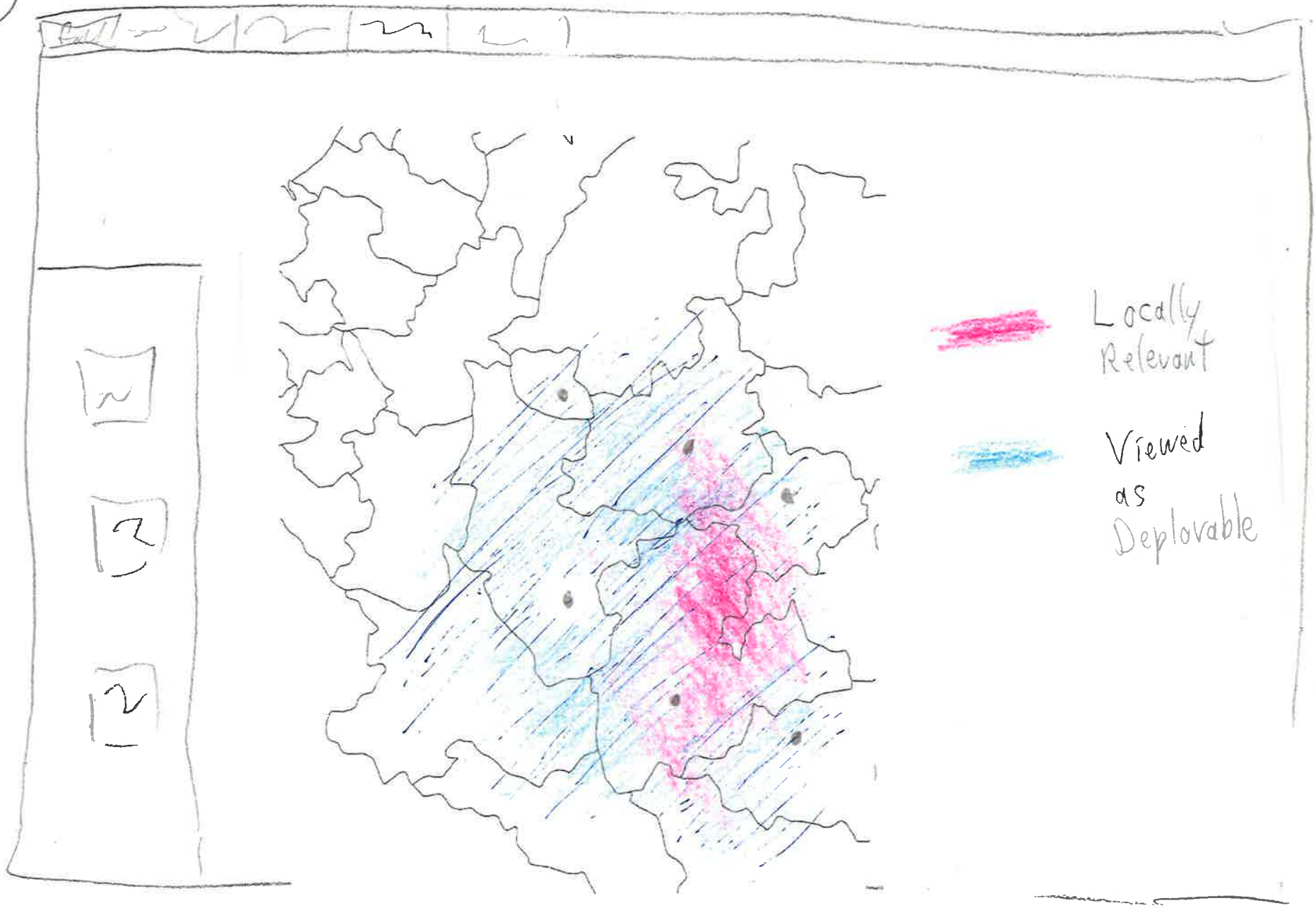
Distant — Local

Just Talk — Call to Action

Old — New

Idiosyncratic — Large-Group

C1



(C2)

Distribution of Stances for Topic Famine or Flood
Region Province Q and within 50 km
Time Aug 30 to now

Bad — Good
Deplorable — Praiseworthy
Subjective — Factual
Typical — Unusual
Distant — Local
Just Talk — Urgent
Old — New
Idiosyncratic — Large-Group

Demographic
mosque-y families
all others

H1

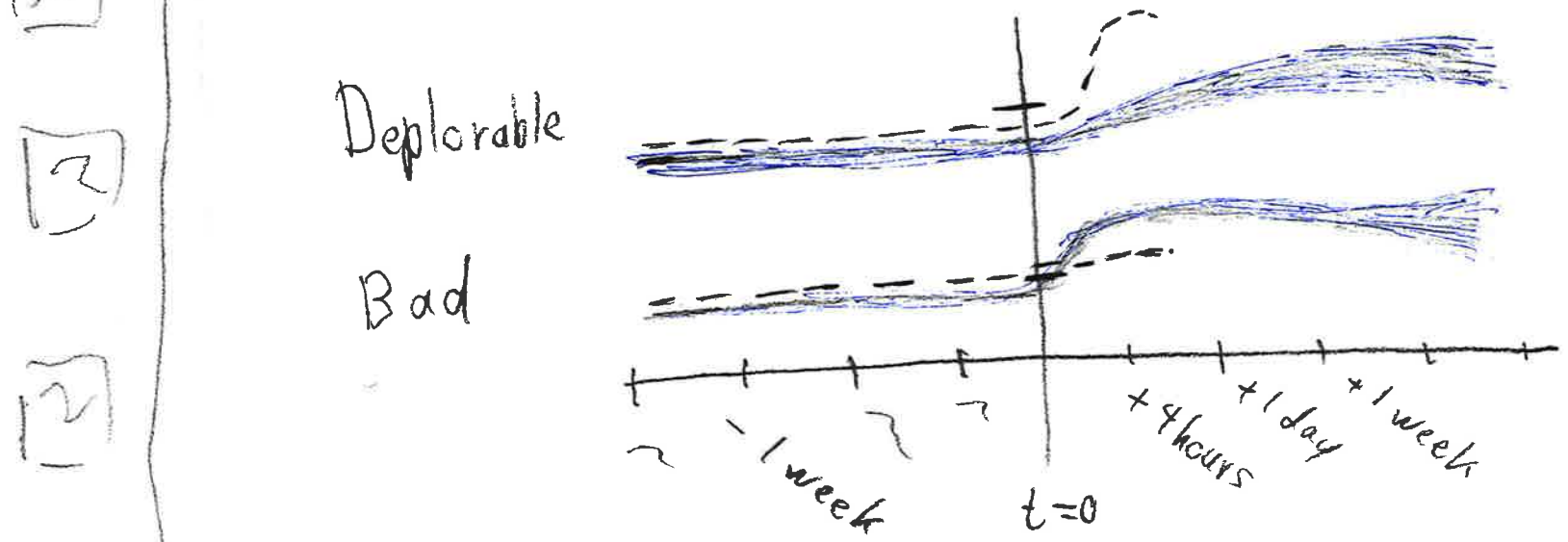
Keyword Density for: Volcanic

H2

Average Stance vs. Time

--- this incident (volcanic, 8-16-2018, Nakuru)

—— typical incident



Stance in Lorelei: Tentative Stance List

1. **Bad Implications … Good Implications**

something with undesirable consequences, such as a raise in taxes, an approaching storm, or a flood …
… something good, such as a peace agreement, a good harvest, or nice weather.

2. **Deplorable Action … Praiseworthy Action**

something viewed as bad attributed to someone or some organization …
… something good attributed to someone or something

3. Subjective … Factual

opinions, of the presenter or someone else, or information reported skeptically or speculatively …
… information presented as fact

4. **Typical or Unsurprising … Unusual or Surprising**

something expected, such as an opposition politician criticizing the government or the stock market fluctuating …
… something quirky, odd, or unexpected

5. **Distant … Local**

something far away …
… something personally relevant to the listening audience, like local weather or close-by rioting

6. Just Talk … Urgent

something that's just background or hypothetical…
… something that may motivate the listening audience to do something soon or now, like take shelter from a storm or vote in today's election

7. Old Information … New Information

a repetition or rehash of something previously reported …
… new information or description of a recent development

8. Idiosyncratic … Relevant to a Large Group

something relevant to just one or a few people …
… something affecting many people

Stance is the speaker's attitude to what he or she is describing, including a very broad set of feelings and speaking styles. This short list is a selection of those most likely to be useful and easily presentable in an interface. More information is at http://www.cs.utep.edu/nigel/stance/.

Version 3, September 19, 2016. University of Texas at El Paso. nigel@utep.edu

Stance in Lorelei: User-Perspective Study

Overall, is stance functionality something that you'd like to have?

Not Useful 1 2 3 4 5 Definitely Useful

Why did you give it this rating?

For each specific stance scale, how useful does it seem?

Bad … Good

Not Useful 1 2 3 4 5 Definitely Useful

Deplorable … Praiseworthy

Not Useful 1 2 3 4 5 Definitely Useful

Subjective … Factual

Not Useful 1 2 3 4 5 Definitely Useful

Typical … Unusual

Not Useful 1 2 3 4 5 Definitely Useful

Distance … Local

Not Useful 1 2 3 4 5 Definitely Useful

Just Talk … Urgent

Not Useful 1 2 3 4 5 Definitely Useful

Old Information … New Information

Not Useful 1 2 3 4 5 Definitely Useful

Idiosyncratic … Relevant to a Large Group

Not Useful 1 2 3 4 5 Definitely Useful

For the most useful ones, please note why you rated them highly.

What additional stance-type scales would you like to have, if any?

What additional scenario types should we be considering?

How could we improve the interface for working with stance?

Are there any of these stances that could be especially useful for some particular type of humanitarian assistance scenario? (e.g. flood, famine, disease, earthquake, hurricane, wildfire, civil unrest, crime, terrorism)

Any other comments? (Things we should know, or should work on, or should suggest to DARPA, etc.)