Proto MT Evaluation for Humanitarian Assistance Disaster Response Scenarios

Doug Jones

Association for Machine Translation in the Americas

31 OCT 2016

This material is based upon work supported by the Defense Advanced Research Projects Agency under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Defense Advanced Research Projects Agency or the U.S. Air Force. Approved for public release; distribution unlimited. Delivered to the U.S. Government with Unlimited Rights, as defined in DFARS Part 252.227-7013 or 7014 (Feb 2014). Notwithstanding any copyright notice, U.S. Government rights in this work are defined by DFARS 252.227-7013 or DFARS 252.227-7014 as detailed above. Use of this work other than as specifically authorized by the U.S. Government may violate any copyrights that exist in this work. © 2016 Massachusetts Institute of Technology.
How might mission impact be measured?

SPOILER ALERT

MT may be neither necessary nor sufficient for HADR mission relevance
Top 20 Languages of the World

Legend and Source

Legend and Source

~7500 Languages
~7500 Languages – 1500 endangered ≈ 6500

≈ 1% MT
Use Standards and Doctrine

Arabic Defense Language Proficiency Test 5

Performance Measures

1. Stopped the vehicle. — —
2. Informed the occupants of the reason for the search. — —
3. Identified the occupants by looking at their drivers' licenses or ID cards. — —
4. Directed the occupants to get out of the vehicle. — —
5. With the exception of the driver, directed the occupants to move to a place about 5 meters from the vehicle and out of the flow of traffic where they could be observed. — —
6. Directed the driver to open all doors and compartments, to include the ashtray, glove box and/or armrest, trunk, and hood. — —
7. Search the vehicle in a sequenced manner. — —

STP 21-1-SMCT 18 June 2009 171-137-0001 3-199
Flow of Information from Needs to Response

Individual Needs

We Need Help

Deliver Assistance

General Mission

Aggregate Needs

Military Tasking
Flow of Information from Needs to Response

Individual Needs

WE NEED FOOD

Deliver Assistance

Actionable Information

General Mission

Aggregate Needs

Military Tasking
LORELEI HADR Topics
Low Resource Languages for Emergent Incidents

Needs
• Water
• Food
• Shelter
• Medical
• Utilities, Energy, Sanitation
• Infrastructure Damage
• Search / Rescue
• Evacuation

Obstacles
• Civil Unrest
• Extreme Violence
• Regime Change
Mission Tasking Matrix (MITAM)

**USAID/OFDA DoD Mission Tasking Matrix (MiTaM)**

**RESPONSE:**
- **TITAN Earthquake** (EXERCISE ONLY)
- New Missions identified as of **1-Oct-11**
- at **1500 Local**

**Mission #**
- **T-101**
- **URGENT**

**WHO**
- **WHO is Requesting US Military Assistance?**
  - **Name, Pos:** Chivers, Dana
  - **Org/Office:** OFDA DART Civ-Mil Coordinator
  - **email:** dchivers@ofda.gov
  - **phone:** +1. 571.594-3937

**WHAT**
- **WHAT type of Service or Goods are Requested?**
  - **Describe as clearly as possible what you want the military to do**
  - Transport DART team on Aerial Recon of effected routes (Hwy1 & 2) to the south of the Capitol

**WHEN**
- **WHEN is it needed?**
  - **Date(s) & Time(s):** ASAP - request NLT 24 hours from now.

**WHERE**
- **WHERE is it needed? ...and HOW**
  - **If the request is for a static position:**
    - **Site Name:** N/A
    - **Grid:** N/A
    - **POC on-site & contact info:** N/A

**WHAT CARGO needs to be moved?**

<table>
<thead>
<tr>
<th>Total # pieces</th>
<th>What</th>
<th>Total Weight (units)</th>
<th>Total Volume (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NONE</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PASSENGERS to be moved?**

- **TOTAL:** 5
  - **Time on Objective & Special Instructions:** Mr. Harnii will need to be dropped off at the village of Queria. All others get off with him for a 30 minute meeting vic the LZ, then need to get back on for transport back to base. Mr Harnii will stay at the village.

**WHY**
- **WHY is this requested of the military?**
  - **Reason:** No civilian assets available at this time
LORELEI Situation Frame Highlights
Low Resource Languages for Emergent Incidents

\[<\text{Situation}> ::= <\text{Need}> \mid <\text{Issue}>\]

\[<\text{Need}> ::= \text{Water Supply} \mid \text{Food Supply} \mid \text{Shelter} \mid \text{Medical Assistance} \mid \text{Utilities, Energy, or Sanitation} \mid \text{Infrastructure} \mid \text{Search/Rescue} \mid \text{Evacuation}\]

\[+ \ (\text{Urgent} \mid \text{Not Urgent}) + (\text{Current} \mid \text{Past} \mid \text{Future})\]

\[+ <\text{Place}>\]

\[<\text{Issue}> ::= \text{Civil Unrest or Widespread Crime} \mid \text{Terrorism or other Extreme Violence} \mid \text{Regime Change}\]

\[<\text{Place}> ::= ( <\text{LOC}> \mid <\text{GPE}> \mid \text{Unknown})\]
LoReHLT16 Evaluations

NIST's new Low Resource human language technologies (LoReHLT) evaluation series aims to advance HLT that can provide rapid and effective response to emerging incidents where the language resources are very limited. Participation in LoreHLT is open to anyone who finds the task(s) of interest. LoReHLT 2016 offers three tasks:

- Machine translation (MT)
- Situation Frame (SF)
- Named entity recognition (NER)

Highlights

- Surprise language evaluation
- Two training conditions: constrained (required) and unconstrained
- Three evaluation checkpoints to gauge performance based on time and training resources given
Data from Haiti 2010 Earthquake

- **Mission 4636**
  - SMS Service during 2010 Earthquake for people to send messages requesting help

- **Ushahidi Collaboration Platform**
  - Volunteers translated and annotated requests and shared with responders

- **Data publicly available**
  - pydata book on github.com
  - Haiti Crisis Map on datahub.io
LORELEI seeks to enable English speakers to respond in HADR scenarios without knowing the local languages.

Sample Mission 4636 Message*

*Notes: (1) character encoding issues in SMS: ŋ/ò; ċ/è; 
(2) cut-off at 160 chars
Nou nan zòn tigwav nou ta renmen pou nou èd men fok se pajan mwen wè mesye minista ap fè moun yo nou bezwen anpil tant ak medikaman pou grip la fyè ak no Time: 2010-01-28 23:31:13

**Situation Frame**

**ID**
HT2010EQ01-U2517

**Type**
Need Shelter, Medical

**Entities**
UNK

**LOC**
Petit Goave

**Time**
2010-01-28 23:31:13

**SEC**
NONE

**Subtopic**
NONE

**From Gazetteer:**

**Name**
Petit Goave

**Equiv**
tigwav

**LAT**
18.37743

**LON**
-72.93157

*Ushahidi Data: https://github.com/pydata/pydata-book/blob/master/ch08/Haiti.csv*
“Signal” in Low Resource Language “Noise”

• Trends and “Dots on a Map”
• Use confidence levels to help aggregate information
Dots on a Map

2a: Food Shortage

Trends in Message Categories

Persons News

Collapsed structure

Water Shortage

Food Shortage
MT may be neither necessary nor sufficient for HADR mission relevance

\[
\text{<Situation> ::= } \text{<Need> | <Issue>}
\]

\[
\text{<Need> ::= Water Supply | Food Supply | Shelter | Medical Assistance | Utilities, Energy, or Sanitation | Infrastructure | Search/Rescue | Evacuation + ( Urgent | Not Urgent ) + ( Current | Past | Future )}
\]

\[
\text{<Issue> ::= Civil Unrest or Widespread Crime | Terrorism or other Extreme Violence | Regime Change}
\]

\[
\text{<Place> ::= ( <LOC> | <GPE> | Unknown )}
\]

High volume, uncertain inputs acceptable if aggregate output is useful
Concept of Evaluation

• Inject LORELEI analytics into HADR exercise from low-resource language messages (social media, i.e.: Twitter)

• Measure outcomes: i.e.:
  – Accuracy of Decision
  – Time to Right Decision
  – Reduction of Overlapping Effort (Efficiency)

• Vary conditions for Situation Frames (SF) from Low Resource Languages: i.e.:
  – Reference SF
  – System SF
Concept of Experiment

Situation Frames

- Reference
- System

Accuracy | Time | Efficiency

Technology Opportunity Zone

HADR Participant Experiments

<Need>
<Issue>
<Place>

Parameters

Situation Frame

Decision
THANKS

Doug Jones  DAJ@LL.MIT.EDU