Proto MT Evaluation for Humanitarian Assistance Disaster Response Scenarios

Doug Jones

Association for Machine Translation in the Americas

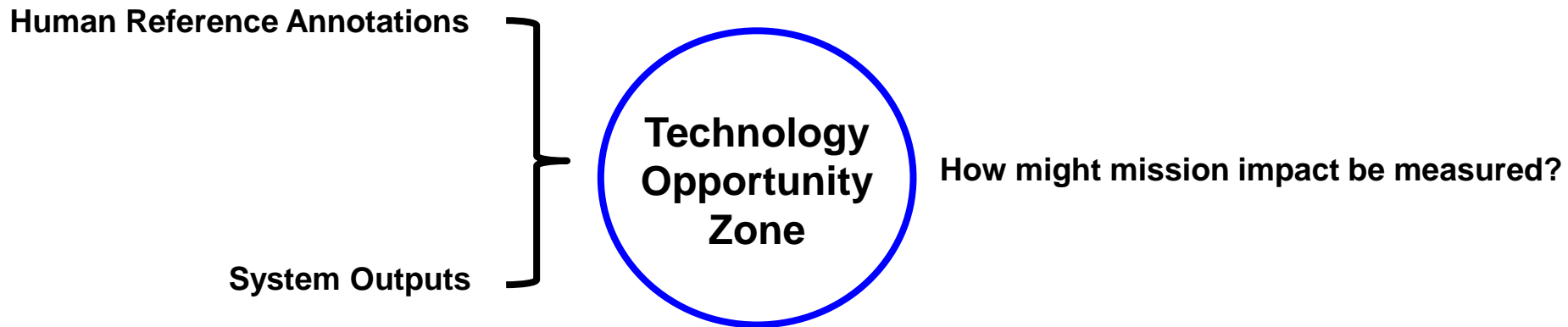**31 OCT 2016**

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Concept of Experiment
(Last Slide)

Human Reference Annotations

System Outputs

Technology Opportunity Zone

How might mission impact be measured?

SPOILER ALERT

MT may be neither <u>necessary</u> nor <u>sufficient</u> for HADR mission relevance

DAJ@LL.MIT.EDU

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

HADR
Humanitarian Assistance Disaster Relief

DAJ@LL.MIT.EDU

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
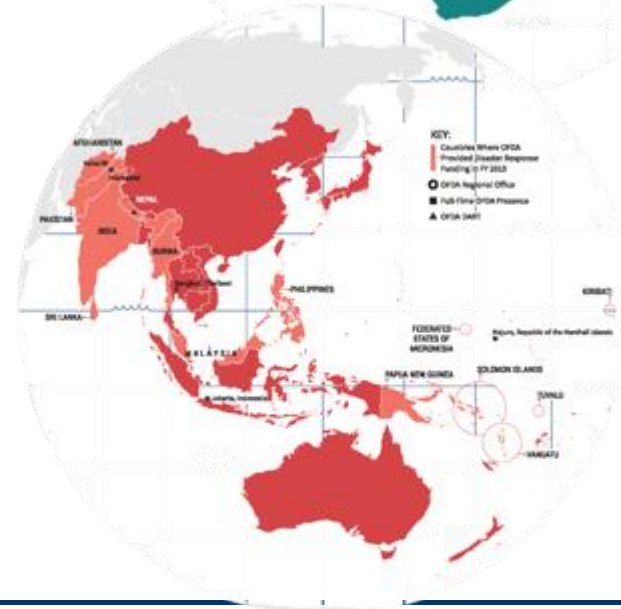
Legend and Source

- Size of circles indicates number of speakers based on Ethnologue data. Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2015. Ethnologue: Languages of the World, Eighteenth edition. Dallas, Texas: SIL International.

Coverage of Seven Available MT Systems

Legend and Source

- Size of circles indicates number of speakers based on Ethnologue data. Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2015. Ethnologue: Languages of the World, Eighteenth edition. Dallas, Texas: SIL International.

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

~7500 Languages

~7500 Languages – 1500 endangered ≈ 6500

≈ 1% MT

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Use Standards and Doctrine

Arabic
Defense Language Proficiency
Test 5

STP 21-1-SMCT

HEADQUARTERS
DEPARTMENT OF THE ARMY

Soldier's Manual of Common Tasks
Warrior Skills
Level 1

JUNE 2009
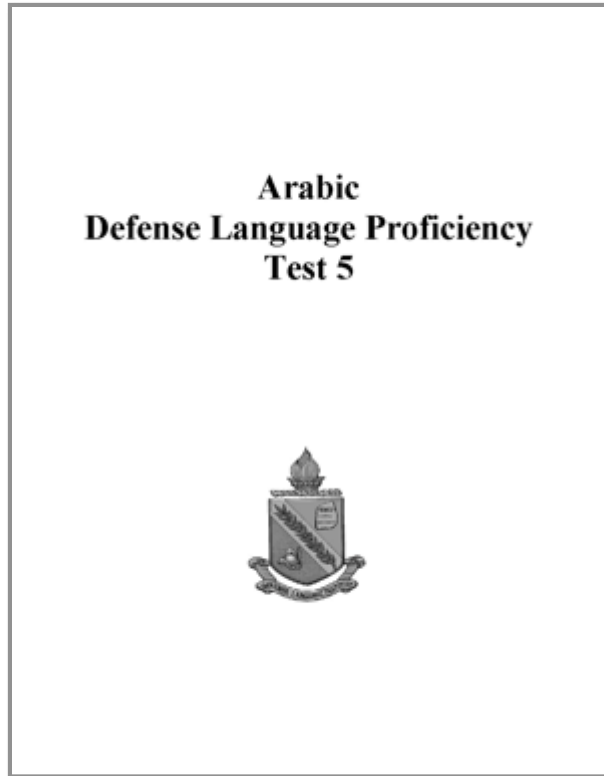
Performance Measures	GO	NO GO
1. Stopped the vehicle.	——	——
2. Informed the occupants of the reason for the search.	——	——
3. Identified the occupants by looking at their drivers' licenses or ID cards.	——	——
4. Directed the occupants to get out of the vehicle.	——	——
5. With the exception of the driver, 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. Searched the vehicle in a sequenced manner.	——	——

STP 21-1-SMCT   **18 June 2009**   **171-137-0001**   **3-199**

DAJ@LL.MIT.EDU

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

15:34

Flow of Information from Needs to Response

Individual Needs

Deliver Assistance

General Mission

Aggregate Needs

Military Tasking

Flow of Information from Needs to Response

Individual Needs

Actionable Information

Deliver Assistance

General Mission

Aggregate Needs

Military Tasking

DAJ@LL.MIT.EDU

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

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**

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

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
Priority	**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?

Total #/ pieces	What	Total Weight ( ) units	Total Volume ( ) units
NONE	NONE		
	total		
Hazmat?			
Special instructions			

PASSENGERS to be moved?   TOTAL: 5

Organization	Name & Position	Nationality
DART	Dana Chivers, Civ-Mil Coordinator	USA
DART	John Gambulputty, DART Tm Ldr	USA
USAID/Titan	Bobby Houne, USAID MDRO	USA
Min of Emerg	Goula Harnii, Vice Minister	Titan
IOM	Francois Dubunet, IOM Field Officer	France

Time on Objective & Special Instructions: Mr. Harnii will need to be dropped off at the village of Queriia. 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?

Is the military your choice of last resort? **No civilian assets availible at this time**

Form:MLU/8007a   rev 16AUG10. Previous editions obsolete

\<Situation\> ::=   \<Need\> ¦ \<Issue\>

\<Need\> ::=   Water Supply ¦ Food Supply ¦ Shelter ¦ Medical Assistance ¦ Utilities, Energy, or Sanitation ¦ Infrastructure ¦ Search/Rescue ¦ Evacuation

+ (Urgent ¦ Not Urgent) + (Current ¦ Past ¦ Future)

+ \<Place\>

\<Issue\> ::=   Civil Unrest or Widespread Crime ¦ Terrorism or other Extreme Violence ¦ Regime Change

\<Place\> ::=   (\<LOC\> ¦ \<GPE\> ¦ Unknown)

LoReHLT16 Evaluations

f 8+ 🐦

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**
  - **http://www.mission4636.org/**

- **Ushahidi Collaboration Platform**
  - **Volunteers translated and annotated requests and shared with responders**
  - **http://www.ushahidi.com/**

- **Data publicly available**
  - **pydata book on github.com**
  - **Haiti Crisis Map on datahub.io**

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

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

LORELEI seeks to enable English speakers to respond in HADR scenarios without knowing the local languages.

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Message #2517[*] from Haiti 2010 Earthquake

> **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:
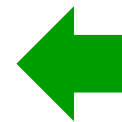
Name	Petit Goave
Equiv	tigwav
LAT	18.37743
LON	-72.93157

*Ushahidi Data: https://github.com/pydata/pydata-book/blob/master/ch08/Haiti.csv

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

"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

- https://github.com/pydata/pydata-book/blob/master/ch08/Haiti.csv
- http://datahub.io/dataset/ushahidi/resource/81d058a8-173a-49d9-8ce9-4edf5e7cafc9

DAJ@LL.MIT.EDU

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
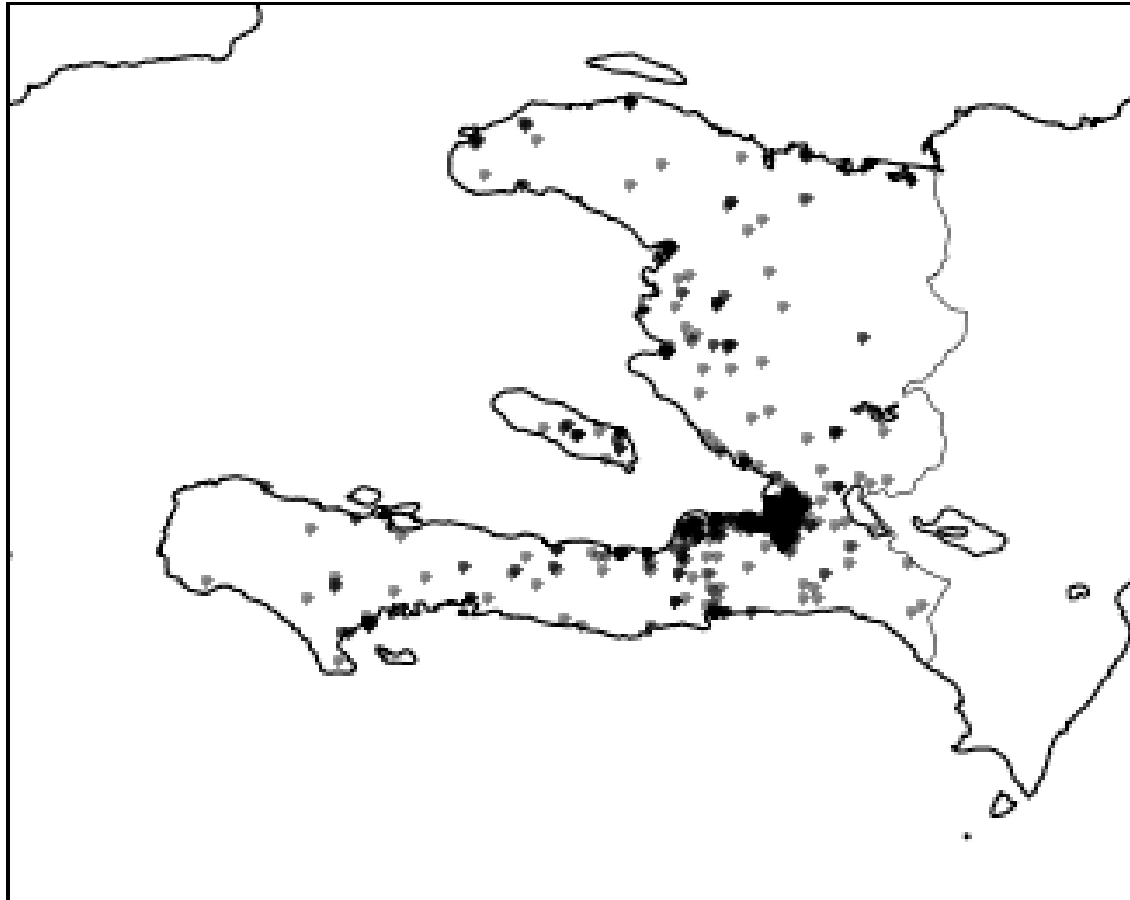
Trends in Message Categories

MT may be neither <u>necessary</u> nor <u>sufficient</u> for HADR mission relevance

Foreign Language →

Foreign Language → **English** →
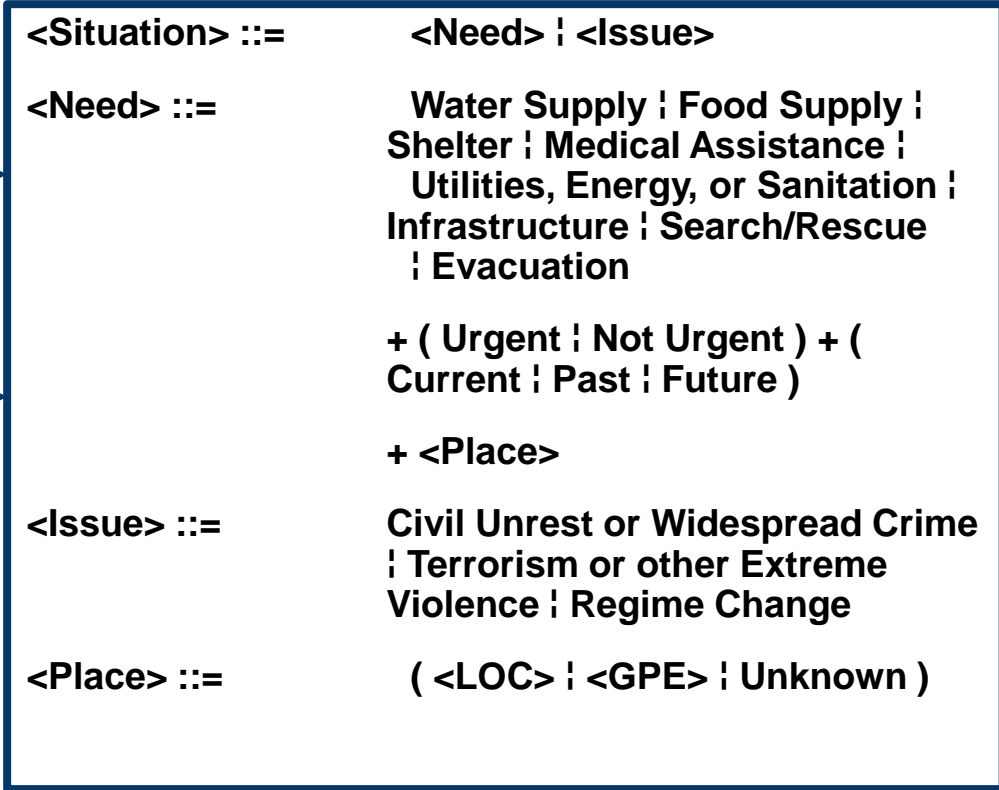
(bracket) **High volume, uncertain inputs acceptable if aggregate output is useful**

```
<Situation> ::=        <Need> ¦ <Issue>

<Need> ::=             Water Supply ¦ Food Supply ¦
                      Shelter ¦ Medical Assistance ¦
                       Utilities, Energy, or Sanitation ¦
                      Infrastructure ¦ Search/Rescue
                       ¦ Evacuation

                      + ( Urgent ¦ Not Urgent ) + (
                      Current ¦ Past ¦ Future )

                      + <Place>

<Issue> ::=           Civil Unrest or Widespread Crime
                      ¦ Terrorism or other Extreme
                      Violence ¦ Regime Change

<Place> ::=            ( <LOC> ¦ <GPE> ¦ Unknown )
```
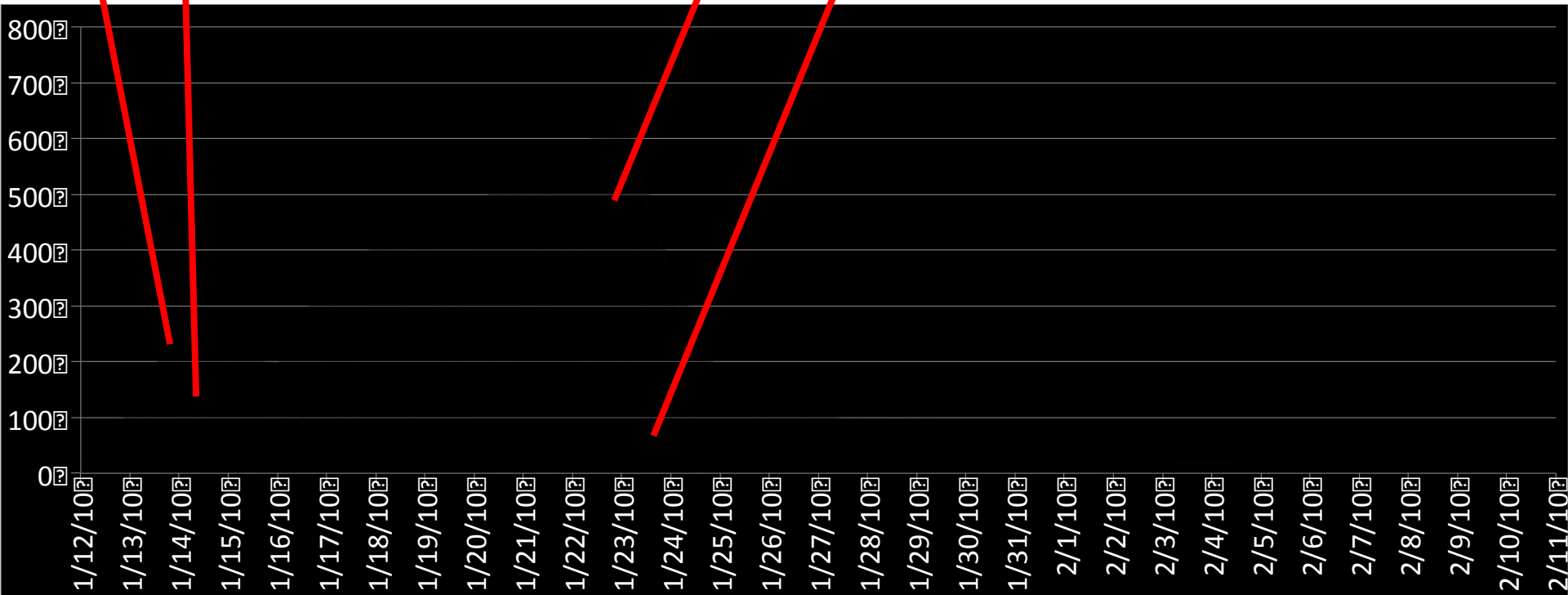
Concept of Evaluation

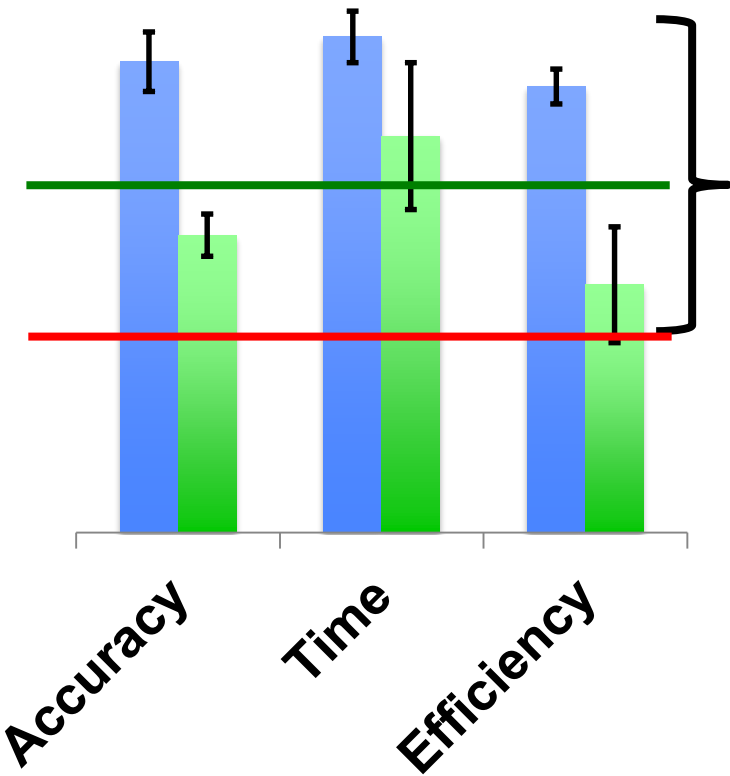- **Inject LORELEI analytics into HADR exercise from low-resource language messages (social media, i.e.: Twitter)**

- **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.:**
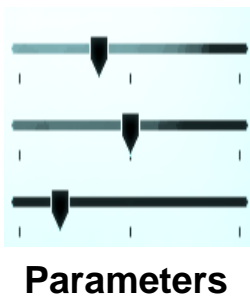 - **Reference SF**
  - **System SF**
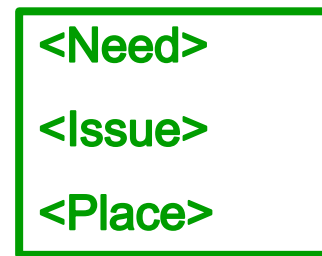
Concept of Experiment
Situation Frames

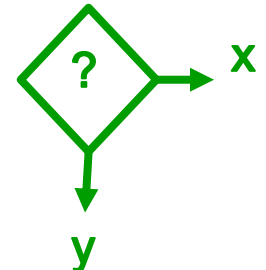**Reference**

System

Accuracy · Time · Efficiency

Technology Opportunity Zone

HADR Participant Experiments

Parameters

<Need>
<Issue>
<Place>

Situation Frame

? → x

y

Decision

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
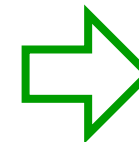
THANKS

Doug Jones DAJ@LL.MIT.EDU

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY