**Slide One**

Overarching Problem

How do we assess the impact and measure the effectiveness of USG government funded media capacity building projects?

Currently Solutions:

MCAT – Media Content Analysis Tool

-Small Samples and Hand coded

-Unrep

There is a better way.

**Slide Two**

Basic Research Questions

Do local media outlets in Liberia adequately provide media coverage of the entire country?

Test Case

How did local media outlets in Liberia cover events before, during and after the Ebola crisis?

Further Questions

What are the predictors of media coverage in Liberia?

Potentially: Geographical, Demographics, Economic Centers, Language, Telecommunications Infrastructure, Politics, other development indicators?

Was media coverage representative of the geographical scope and extent of the Ebola outbreak?

Slide Three

Hypothesis

I hypothesize that timely and accurate media coverage of significant events in Liberia is determined by population density, availability (and functionality) of telecommunications infrastructure, levels of literacy, and ethnic and linguistic divisions.

Slide Four

Data Collection Plan

Scrape all the online media content “stories” from 9 local media outlets.

“The Inquirer”, “The Analyst”, “Front Page Africa”, “The New Dawn”, “Liberian News Agency”, “Post 1847”, “GNN Liberia”, “Daily Observer”, “Government Ministry of Information”

DHS Data – Already Cleaned

**The Demographic and Health Surveys (DHS) Program**has collected, analyzed, and disseminated accurate and representative data on population, health, HIV, and nutrition through more than 300 surveys in over 90 countries.

Ebola Time Series by Region Data from HDX – Already Cleaned

The goal of the Humanitarian Data Exchange (HDX) is to make humanitarian data easy to find and use for analysis.

Slide Five

Media Content

*Data Cleaning*

*- Title*

*- Post Date*

*- Author*

*- Story Lead (First ~4 sentences)*

*- Entire Story*

*Feature Extraction*

*- Named Entity Recognition (Names, Locations, Organizations, etc)*

*- Story Category*

*- Keywords*

*- Sentiment*

*- LDA*

Slide Six

Plan for Modeling

What are the predictors of news coverage? Media content features and DHS data.

With this model-

Look at stories published by day and do a regression of predicted media coverage. As the Ebola outbreak starts, does the near daily casualty data alter news coverage?

When does media coverage shift to providing more or less coverage of a geographical area?

Can we model sentiment of media outlet, author to a geographical area?

Slide Seven

Status

Web scraper Issues:

False Positives in Collection Processes (Redirects from a page that does not exist to the homepage. No 404 error)

Feature Extraction from Media Content

In progress

Need to refine and customize for each page