Group Charlie
Multi-Lingual SMS

Andrew, Chihang, Daria, Swaraj and Tanvi

University of Cambridge

March 4, 2015

Background

- Africa's Voices: Opinion polling done by broadcasting a variety of questions from several radio stations, where answers were received as SMS messages.

Background

- Africa's Voices: Opinion polling done by broadcasting a variety of questions from several radio stations, where answers were received as SMS messages.
- Main problem: The SMS messages were in different languages and sometimes included irrelevant information, such as details of the sender.

Background

- Africa's Voices: Opinion polling done by broadcasting a variety of questions from several radio stations, where answers were received as SMS messages.
- Main problem: The SMS messages were in different languages and sometimes included irrelevant information, such as details of the sender.
- Sub-problem: Machine translation is not feasible for minority languages with little economic value. As a result of this, translation was taken to be an expensive task.

Background

- Africa's Voices: Opinion polling done by broadcasting a variety of questions from several radio stations, where answers were received as SMS messages.
- Main problem: The SMS messages were in different languages and sometimes included irrelevant information, such as details of the sender.
- Sub-problem: Machine translation is not feasible for minority languages with little economic value. As a result of this, translation was taken to be an expensive task.
- Problem we're trying to solve: allow for easy detection of opinion trends in such data, despite it being in several languages.

Our goals

- To create a visual browser that will allow international policy teams to observe trends.

Our goals

- To create a visual browser that will allow international policy teams to observe trends.
- This is a tool for scientists — it should have a friendly and interactive UI while not getting in the way.

Our goals

- To create a visual browser that will allow international policy teams to observe trends.
- This is a tool for scientists — it should have a friendly and interactive UI while not getting in the way.
- Due to translation being expensive, we need to allow for the researchers to determine more important words/topics for which they can prioritise translation.

Our goals

- To create a visual browser that will allow international policy teams to observe trends.
- This is a tool for scientists — it should have a friendly and interactive UI while not getting in the way.
- Due to translation being expensive, we need to allow for the researchers to determine more important words/topics for which they can prioritise translation.

Additional requirements set along the way:

Our goals

- To create a visual browser that will allow international policy teams to observe trends.
- This is a tool for scientists — it should have a friendly and interactive UI while not getting in the way.
- Due to translation being expensive, we need to allow for the researchers to determine more important words/topics for which they can prioritise translation.

Additional requirements set along the way:

- Allow for identification of the language that a message was written in.

Our goals

- To create a visual browser that will allow international policy teams to observe trends.
- This is a tool for scientists — it should have a friendly and interactive UI while not getting in the way.
- Due to translation being expensive, we need to allow for the researchers to determine more important words/topics for which they can prioritise translation.

Additional requirements set along the way:

- Allow for identification of the language that a message was written in.
- Allow for 'cleaning up' of the data by removing names, `txt-speak` and other slang.

Overview of the project

The project has been implemented as a web application.

Overview of the project

The project has been implemented as a web application.
User input:

Overview of the project

The project has been implemented as a web application.
User input:

- CSV file blah blah

Overview of the project

The project has been implemented as a web application.
User input:

- CSV file blah blah
- Categories + example

Overview of the project

The project has been implemented as a web application.
User input:

- CSV file blah blah
- Categories + example

Viewing ‘modes’ implemented:

Overview of the project

The project has been implemented as a web application.
User input:

- CSV file blah blah
- Categories + example

Viewing 'modes' implemented:

- Word cloud

Overview of the project

The project has been implemented as a web application.
User input:

- CSV file blah blah
- Categories + example

Viewing 'modes' implemented:

- Word cloud
- Pie chart

Overview of the project

The project has been implemented as a web application.
User input:

- CSV file blah blah
- Categories + example

Viewing 'modes' implemented:

- Word cloud
- Pie chart
- Bar graph

Overview of the project

The project has been implemented as a web application.
User input:

- CSV file blah blah
- Categories + example

Viewing 'modes' implemented:

- Word cloud
- Pie chart
- Bar graph
- Table view

Overview of the project

The project has been implemented as a web application.
User input:

- CSV file blah blah
- Categories + example

Viewing 'modes' implemented:

- Word cloud
- Pie chart
- Bar graph
- Table view

A data manipulation page has also been added.

Word Cloud

Pie Chart

Bar Graph

Data Manipulation

Technical details

Lessons learned