VAST 2021 Mini Challenge 3 – The Kronos Incident

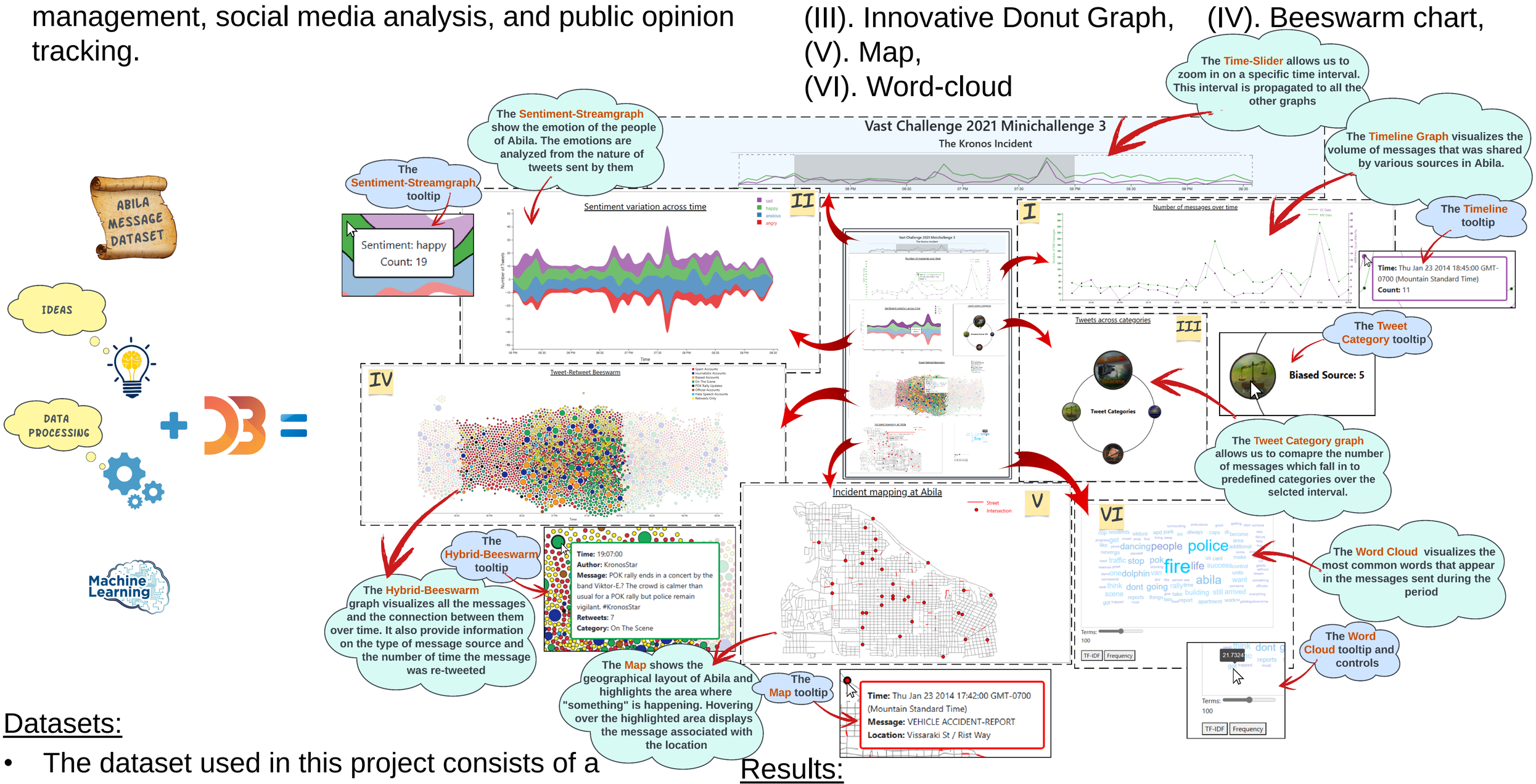
Ira A. Fulton Schools of Engineering
Arizona State University

Introduction

- The VAST 2021 Mini-Challenge 3 aims at analyzing microblogging data related to the events on the fictional island of Abila.
- Abila on Kronos island is GAStech's gas production site.
- GAStech's successful IPO in Jan 2014 was marred by the disappearance of several employees, allegedly due to the involvement of the POK.
- Following visualizations try to analyze microblogs and emergency call data to evaluate risk levels. They aim to identify relevant microblog records related to the crisis.
- Approach and visualizations used here have potential for use in real-world scenarios such as crisis management, social media analysis, and public opinion tracking.

System Description

- The project incorporated six visualizations and a slider function to enable users to focus on a specific time period.
- The project utilized machine learning techniques to analyze the emotions conveyed by the individuals who posted the tweets.
- Python was used to preprocess the data for certain graphs, whereas JavaScript was used to process the data on page load for the remaining plots.
- All visualizations were developed using the D3 v7 JavaScript library.
- The following visualizations are created using the dataset:
  (I). Timeline Graph (II). Stream Graph,
  (III). Innovative Donut Graph, (IV). Beeswarm chart,
  (V). Map,
  (VI). Word-cloud



Datasets:

- The dataset used in this project consists of a collection of static microblog records obtained during the crisis at Abila, organized in a tabular format with multiple columns such as type, date and time, author, message, and location.
- The data is stored as csv files, which are commonly used file formats for tabular data.
- The dataset comprises two distinct types of data - ccdata and mbdata. The ccdata includes text transcripts of emergency dispatches, while the mbdata comprises microblog records that are similar in nature to Twitter data and are authored by users.
- Additionally, the dataset includes geospatial data in the form of .shp and .dbf files, which were transformed into geojson data using a conversion process to make them compatible with D3.
- By utilizing the geojson data, it was possible to generate a map of Abila and locate specific places by name on the map.
- To prepare the microblog data for visualization purposes, a set of Python scripts were utilized. These scripts implemented various data manipulation techniques to transform the raw data into a format that was suitable for creating the desired visualizations.

Results:

- Complex visualization techniques, such as timeline slider, beeswarm and geographic mapping helped us represent our data in unique ways, giving new insights.
- Preprocessing and cleaning our raw data into a more manageable format saved time and effort.
- We learned to choose and design appropriate visualizations for the nature of our data to provide actionable information to our audience.
- Working with kml and dbf file formats to plot geographical markings presented challenges but converting kml to geojson format enabled accurate plotting on our visualizations.
- Using the created visualization, we were able to :
  - Characterize different content types in the dataset using visual analytics and distinguish meaningful event reports from typical chatter, junk or spam.
  - Use visual analytics to represent and evaluate the evolution of public risk level over the course of the evening.
  - Determine the most critical place to send a team of first responders by providing rationale based on analysis.
- Proposed improvements for the project include optimizing computation, enhancing interactivity, handling dynamic data, and implementing a better responsive design.

Aakash Mathai
amathai1@asu.edu

Dhruv Vohra
dvohra1@asu.edu

Syed Asad Husain
shusain6@asu.edu

Hanish Rao Saini
hsaini4@asu.edu

Sathwik Katakam
skataka@asu.edu

Karthik Aravapalli
karavapa@asu.edu