

THE USE OF SOCIAL MEDIA BY THE ARMY

Ryan Whitcomb, Emma Baldwin, and Ian McCormack

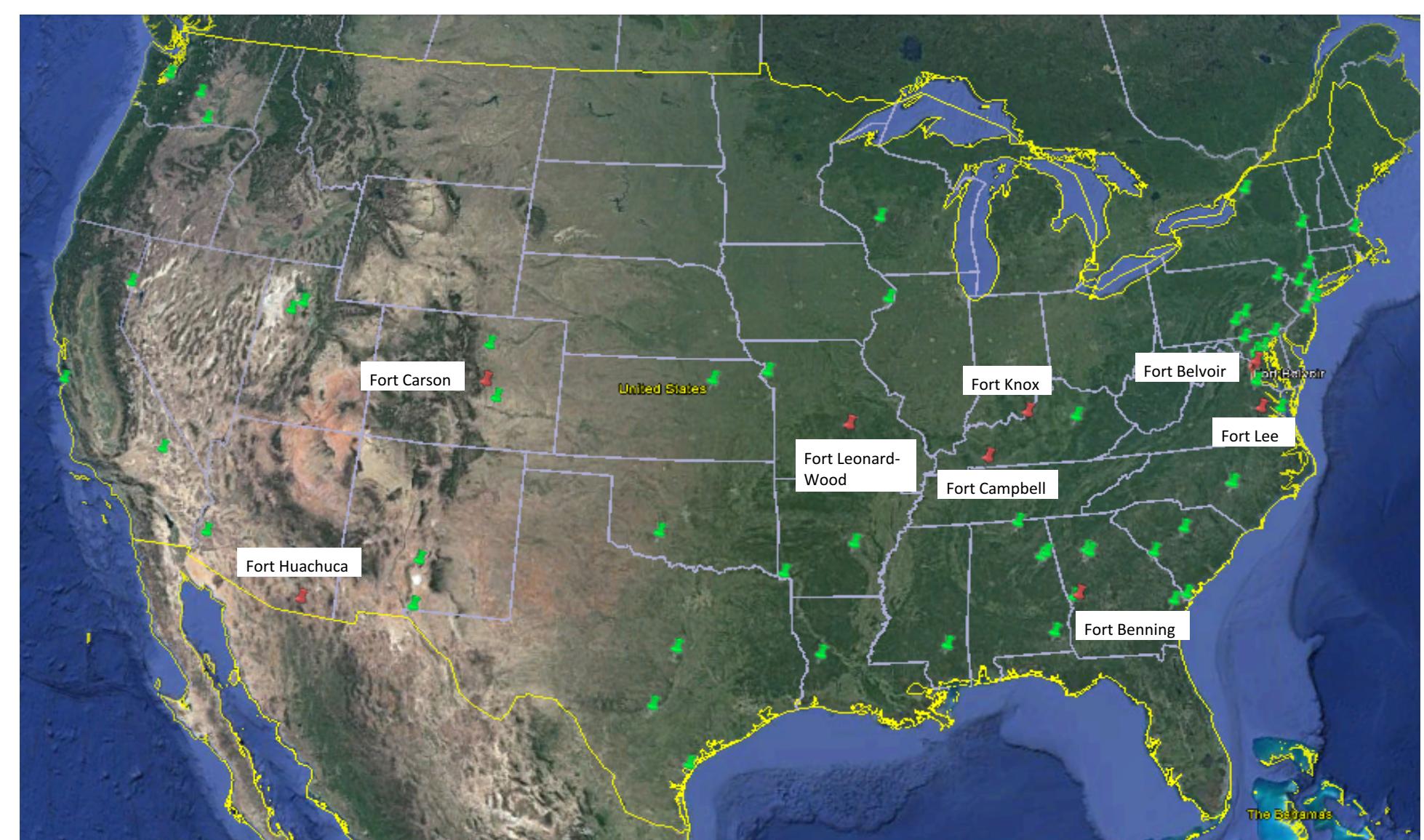
In Collaboration with Dr. Gizem Korkmaz

The Social and Decision Analytics Laboratory, Biocomplexity Institute of Virginia Tech

Department of Computational Modeling & Data Analytics, Virginia Tech

Research Question

Social media offers a new opportunity to capture soldiers' well-being and concerns about the Army. This exploratory study aims to understand the use of Facebook by the U.S. Army population. Data will be captured and used from Facebook groups to determine whether there are differences in topics discussed and sentiments across groups (e.g., soldiers vs. family) and Army bases.



Locations of the Army Bases analyzed (marked in red)

Data & Methods

- As shown in the table below, Facebook has one of the highest concentration of active army news and posts. It is data heavy by focusing on Army specific pages, which allows for large amounts of useful data.

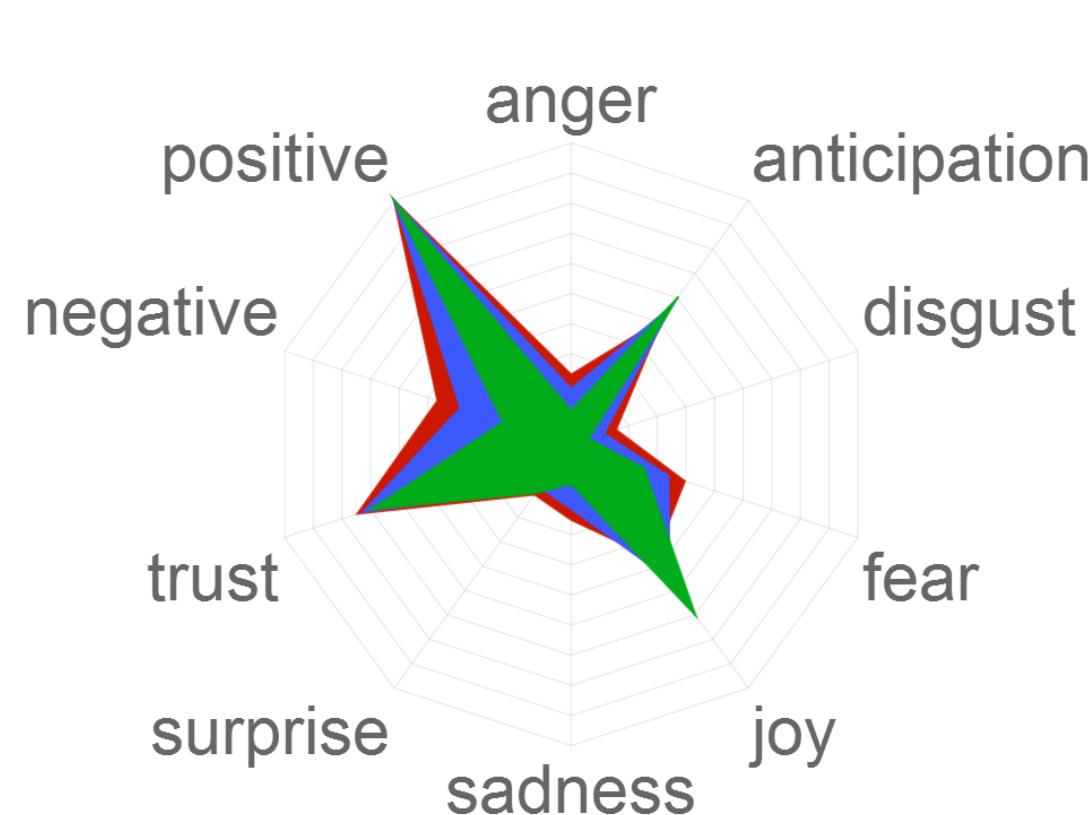
Design Matrix showing the process of selecting which social media to use

	Facebook	Twitter	Instagram	YouTube	Google+	Reddit	LinkedIn	Tumblr	Blogs
Army Population	0.30	5	1.50	5	1.50	4	1.20	5	1.50
Unidentifiable	0.20	2	0.40	3	0.60	3	0.60	2	0.40
Data Heavy	0.16	5	0.80	5	0.80	3	0.48	2	0.32
Searchable	0.14	5	0.70	3	0.42	3	0.42	1	0.14
Easily to Scrape	0.12	5	0.60	5	0.60	2	0.24	5	0.70
Easily Verified	0.08	4	0.16	1	0.08	3	0.24	4	0.16
Score		4.16		4		3.48		2.58	

- Scraping data from Facebook is done through the Facebook API which contains three main data elements to access the desired data:
 - Nodes - The main elements of Facebook, e.g. a User, a Photo, a Page, a Comment
 - Edges - The link between these elements such as a Page's photo or a Photo's comments
 - Fields - Information about an element such as a Person's birthday or the name of a Page
- The Facebook API is accessed using the Python SDK for Facebook's Graph.
- Storing the large amount of data in the project utilizes a MySQL database to safely store and clean the data.
- Data scraping is completed in parallel on the Advanced Research Computing (ARC) Cluster at Virginia Tech.

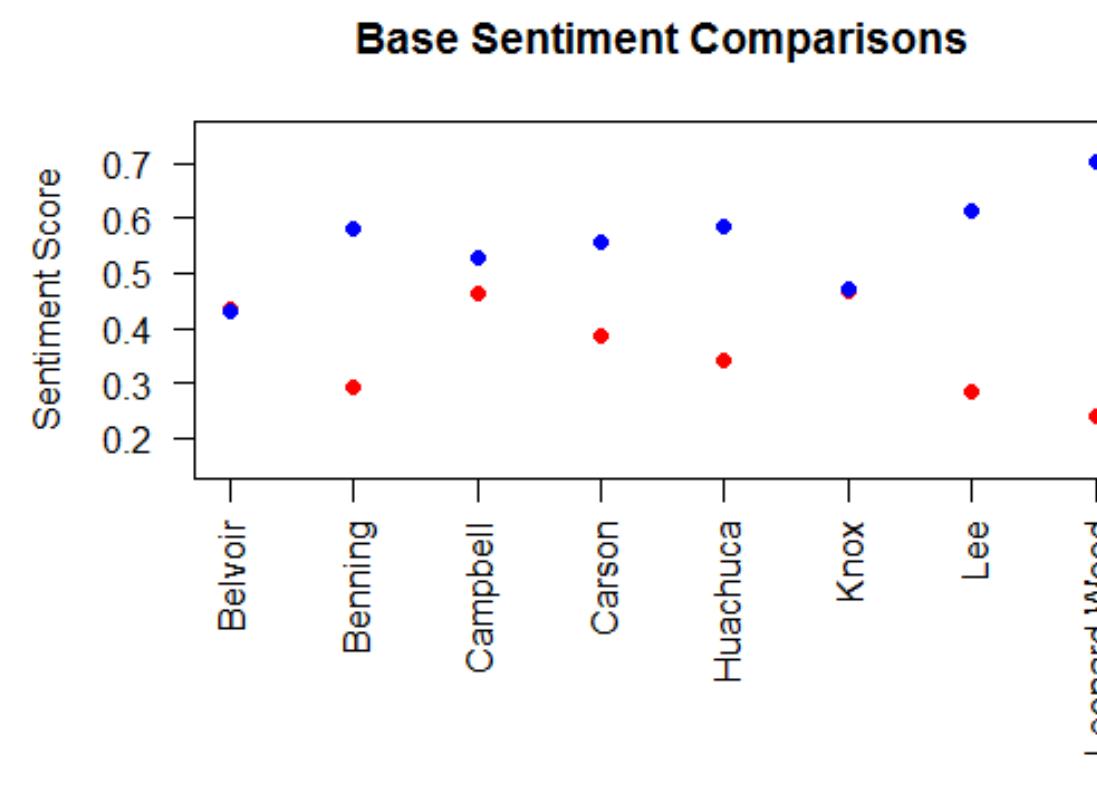
Analysis

Sentiment Distribution

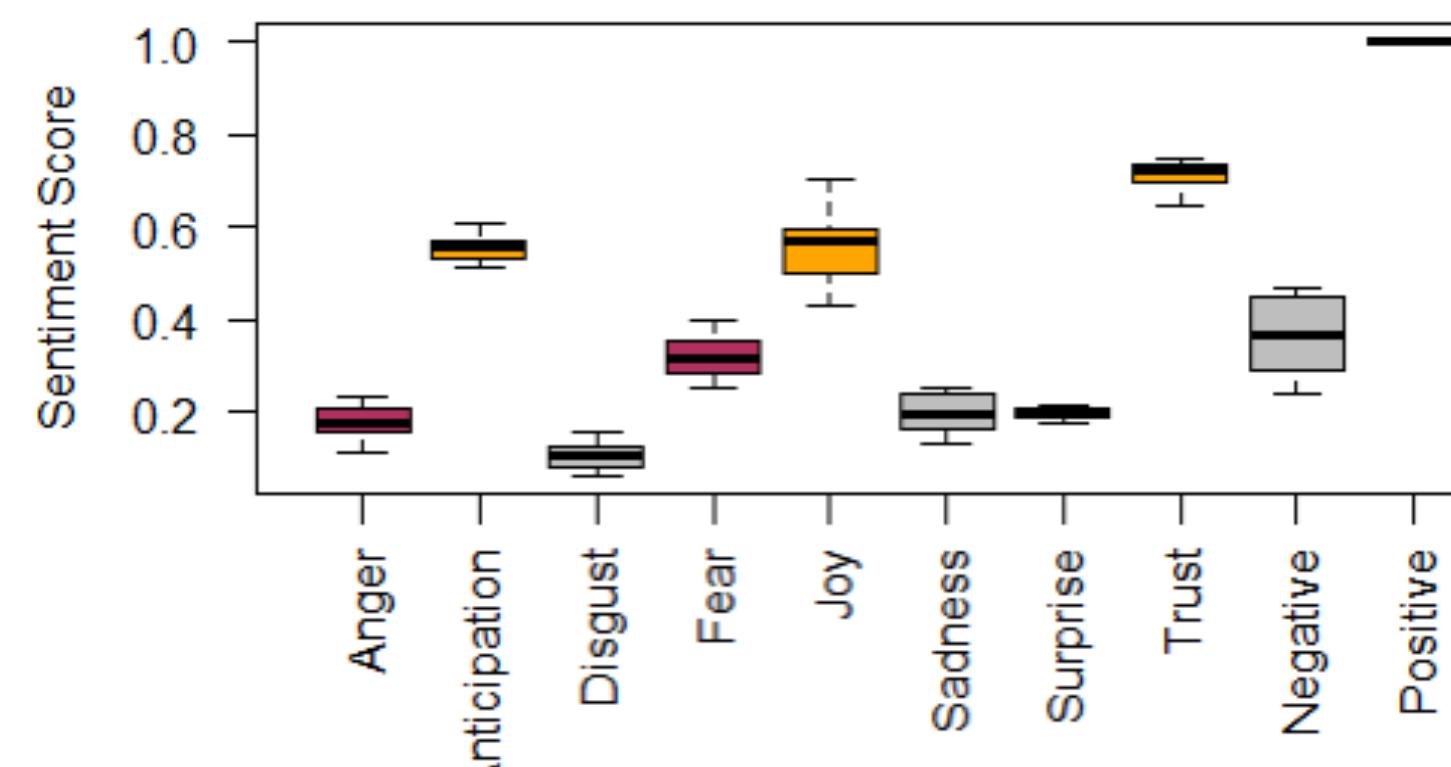


- Facebook sentiment is largely positive with some peaks in anticipation and trust.
- Three different bases, **Fort Knox**, **Fort Carson**, and **Fort Wood**, are shown to demonstrate the differences between different Army populations.

- The two most varied sentiment metrics, **Anger** and **Joy**, show the wide variety reported across the bases.
- Negative sentiment and joy sentiment in the bases tend to be negatively correlated.



Variability of Base Sentiment



Boxplots of the different sentiment metrics that are measured for all bases show an insight into which sentiment varies the most.

- Negative sentiment in the comments has the highest variation, while the positive sentiment has the lowest.

Future Work

- After the data are gathered, the main goal will be to develop models and to use novel visualization techniques to explore sentiment of Facebook posts and comments and to analyze how they are affected by the policy changes over time.
- We plan to extend the analysis to all 83 Army bases and study whether there are differences in sentiments, and how these correlate with attrition rates for those bases.